THE DANIELL APPROACH TO INTEGRATION AND MEASURE

A THESIS

Presented to

The Faculty of the Graduate Division

by

Richard Maurice Ingle

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Applied Mathematics

Georgia Institute of Technology

January, 1969

THE DANIELL APPROACH TO INTEGRATION AND MEASURE

Approved:

Chairman

Date approved by Chairman: February 12, 1969

ACKNOWLEDGMENTS

I wish to express my sincere gratitude to Dr. Eric R. Immel, my thesis advisor, for his expert guidance in the preparation of this thesis. I also wish to thank the other two members of my committee, Dr. James W. Walker and Dr. Harold A. Gersch, for their careful reading of the manuscript. Finally, special thanks are due Mrs. Betty Sims for her excellent typing of the thesis.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

CHAPTER I

INTRODUCTION

The standard measure-theoretic approach to integration theory involves the consideration of a σ-algebra of subsets of a given space X, and the establishment of a certain type of set function called a measure defined on the σ-algebra. The integral is then defined relative to, and in terms of, the measure. Prior to 1917, it occurred to P. J. Daniell that a more direct approach to integration theory results if the integral is defined as a continuous, positive linear functional on a vector lattice. In 1953, L. H. Loomis gave an outline of this approach to integration in *An Introduction to Abstract Harmonic Analysis* [3]. The goal of this thesis is to present a complete and reasonably self-contained development of basic integration and measure theory utilizing the linear functional approach. The work is based, primarily, on the above-mentioned rather brief treatment due to Loomis. A knowledge of the measure-theoretic approach to integration is assumed, and a sufficient reference for the purpose is *Real and Complex Analysis* by W. Rudin [5]. The notation and terminology will, whenever possible, be close to that used by Rudin.

The Daniell integral is developed in Chapter II by means of stages, involving an initial definition followed by an extension procedure. As a part of the extension procedure, such theorems as

Lebesgue's monotone convergence theorem and Lebesgue's dominated convergence theorem are proved, and used in the sequel.

A short discussion of L^p-spaces is contained in Chapter III. Since proofs of the basic theorems concerning L^p-spaces are virtually identical for this approach to the integral with those for the measure-theoretic approach, the theorems are offered without proof. The linear-functional version of the Radon-Nikodym theorem is stated and proved. The method of proof is due to J. von Neumann.

The remaining chapters are actually specializations of the theory of various kinds, useful in applications. An interesting aspect of the Daniell approach is that the integral is attained without any reference to measure theory, but once the integral is attained, a positive measure can be determined in terms of the integral. (This constitutes a direct reversal of the process used in the measure-theoretic approach to integration.) In Chapter IV it is proved that the initial vector lattice and the initial Daniell integral can be chosen in such a way that the measure induced by the resulting integral is the classical Lebesgue measure in Euclidean spaces R^k. The work here is based partly on a discussion in A. E. Taylor, *General Theory of Functions and Integration* [8]. In the final chapter, we restrict attention to integrals that result when the initial vector lattice is the class of continuous, real-valued functions of compact support defined on a locally compact Hausdorff space. The principal result is a proof of the Fubini theorem. Some particular examples relevant to the Fubini theorem are also included in this chapter.

CHAPTER II

THE DANIELL INTEGRAL

Section 1: Definition of Integral

Suppose that L is a vector space (over the real field) of bounded, real-valued functions defined on a set X, and that L is closed under the formation of $f \vee g = \max(f,g)$ and $f \wedge g = \min(f,g)$. Such a space will be referred to as a *vector lattice* of functions on X. Note that if we define $f^+ = f \vee 0$ and $f^- = -(f \wedge 0)$ for any $f \in L$, then $f^+ \in L$ and $f^- \in L$. Also, if $f \in L$ then $|f| \in L$, since $|f| = f^+ + f^-$.

2.1 Definition: Any real-valued functional I on L which has the following three properties is called an *integral* on L.

(1) I is linear: $I(\alpha f+\beta g) = \alpha I(f) + \beta I(g)$ for any $f,g \in L$ and any real numbers α,β.

(2) I is positive: $I(f) \geq 0$ whenever $f \in L$ and $f \geq 0$.

(3) I is continuous under monotone limits: $I(f_n) \downarrow 0$ whenever $f_n \downarrow 0$ and each $f_n \in L$.

Note: "$f_n \downarrow 0$" means that the sequence $\{f_n(x)\}$ is (pointwise) monotone decreasing to zero for each $x \in X$. The positivity property (2) and linearity imply the following additional property:

(2') If $f \geq g$ and $f,g \in L$, then $I(f) \geq I(g)$.

Proof: By property (1), $I(f-g) = I(f) - I(g)$, and by property (2) $I(f-g) \geq 0$, since our hypothesis is that $f - g \geq 0$. Hence $I(f) - I(g) \geq 0$, and thus $I(f) \geq I(g)$.

Section 2: Extension of the Integral

We now wish to extend I to a larger class of functions having all the properties of L, and which is also closed under the formation of monotone limits of its members. As an example, we can let X be the closed interval $[0,1]$, let L be the class of real-valued continuous functions on $[0,1]$, and let I be the ordinary Riemann integral. Properties (1) and (2) are well known, and (3) follows from the fact that on a compact space pointwise monotone convergence to zero implies uniform convergence (Dini's theorem). The extension of L will then be the class of Lebesgue integrable functions on $[0,1]$, and the extended functional I will be the ordinary Lebesgue integral. (This example is a motivation for the detailed work done in Chapter IV.)

We begin by extending I to functions which are limits of monotone increasing sequences in L. Let U be the class of limits of all monotone increasing sequences of functions in L. Here, we include $+\infty$ as a value of the limit function so that any monotone increasing sequence of real-valued functions is considered as convergent. For any $f \varepsilon L$, the sequence $\{f_n\}$ with $f_n = f$, $n = 1,2,\ldots$, is a monotone increasing sequence in L with limit f. Hence $f \varepsilon U$. We conclude that $L \subset U$. Clearly U is closed under addition, multiplication by non-negative constants, and the lattice operations leading to $f \vee g$ and $f \wedge g$.

We now extend I from L to all of U by defining:

$$I(f) = \lim_{n \to \infty} I(f_n)$$

where each $f_n \varepsilon L$ and $f_n \uparrow f \varepsilon U$, and where $+\infty$ is allowed as a possible value for $I(f)$. It will follow from Lemma 2.2 that this definition of I is independent of the particular sequence $\{f_n\}$ converging to f. Thus it is evident that this new definition leaves I as it was on L (if $f \varepsilon L$, put each $f_n = f$). Also, the extended functional I satisfies (1) with non-negative scalars. It will also follow from Lemma 2.2 that I satisfies (2').

2.2 Lemma: If $\{f_n\}$ and $\{g_m\}$ are increasing sequences of functions in L such that $f_n \uparrow f$ and $g_m \uparrow g$, and if $f \leq g$, then $\lim_{n\to\infty} I(f_n) \leq \lim_{m\to\infty} I(g_m)$.

Proof: Fix n, and consider the monotone decreasing (non-increasing) sequence $\{h_m\}$ of functions in L given by $h_m = f_n - g_m$, $m = 1,2,\ldots.$ Now $\lim_{m\to\infty} h_m = f_n - g \leq f - g \leq 0$, since $f_n \leq f \leq g$. We put

$$h_m^+(x) = \begin{cases} h_m(x) & \text{if } h_m(x)>0 \\ 0 & \text{if } h_m(x)\leq 0 \end{cases}$$

and note that $h_m^+ \downarrow 0$ since $f_n - g_m \leq g - g_m$ and $g_m \uparrow g$. We may now apply property (3) to the sequence $\{h_m^+\}$ to conclude that $I(h_m^+)\downarrow 0$. Also, since $h_m \leq h_m^+$ we have, by property (2'), that $I(h_m) \leq I(h_m^+)$ for each $m = 1,2,\ldots.$ We conclude then that

$$\lim_{m\to\infty} I(h_m) \leq \lim_{m\to\infty} I(h_m^+) = 0.$$

That is, $\lim_{m\to\infty} [I(f_n)-I(g_m)] \leq 0$ or $I(f_n) \leq \lim_{m\to\infty} I(g_m)$. But this can be done for each $n = 1,2,\ldots$. Consequently

$$\lim_{n\to\infty} I(f_n) \leq \lim_{m\to\infty} I(g_m)$$

and the proof is complete.

We now show that the extension of I to U by way of $I(f) = \lim_{n\to\infty} I(f_n)$ where $f_n \uparrow f$, $f_n \varepsilon L$, is uniquely defined. Let $f \varepsilon U$ and let $f_n \uparrow f$ and $g_m \uparrow f$ where $\{f_n\}$ and $\{g_m\}$ are sequences in L. Since $f \leq f$ we can apply Lemma 2.2 once to conclude that

$$\lim_{n\to\infty} I(f_n) \leq \lim_{m\to\infty} I(g_m),$$

and again to conclude that

$$\lim_{m\to\infty} I(g_m) \leq \lim_{n\to\infty} I(f_n).$$

Hence we have

$$\lim_{m\to\infty} I(g_m) = \lim_{n\to\infty} I(f_n) = I(f).$$

This shows that our extension is uniquely defined.

This extension of I to U satisfies property (2') on all of U for let $f \varepsilon U$ and $g \varepsilon U$ with $f \leq g$. Then there exist sequences $\{f_n\}$ and $\{g_m\}$ in L such that $f_n \uparrow f$ and $g_m \uparrow g$, and by Lemma 2.2

$$\lim_{n\to\infty} I(f_n) \leq \lim_{m\to\infty} I(g_m).$$

Consequently, $I(f) \leq I(g)$.

2.3 Lemma: If $f_n \varepsilon U$, $n = 1,2,\ldots$, and $f_n \uparrow f$, then $f \varepsilon U$ and $I(f_n)\uparrow I(f)$.

Proof: For each function f_n we construct an appropriate defining sequence of functions in L:

$$\begin{array}{l} f_{11} \leq f_{12} \leq \cdots < f_{1j} \leq \cdots, \quad f_{1j}\uparrow f_1 \\ f_{21} \leq f_{22} \leq \cdots \leq f_{2j} \leq \cdots, \quad f_{2j}\uparrow f_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ f_{n1} \leq f_{n2} \leq \cdots \leq f_{nj} \leq \cdots, \quad f_{nj}\uparrow f_n \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \end{array}$$

Now let $g_n = f_{1n} \vee f_{2n} \vee \cdots \vee f_{nn}$, for each positive integer n. (Note that if the above array of functions is viewed as an infinite-dimensional matrix, then g_n is the lattice maximum of the first n functions in the nth column.) By our initial closure assumptions regarding L, each $g_n \varepsilon L$. Clearly $\{g_n\}$ is an increasing sequence. Let k be an arbitrary positive integer. Fix k. Now for each $n \geq k$ we have

$$g_n = f_{1n} \vee \cdots \vee f_{kn} \vee \cdots \vee f_{nn} \geq f_{kn}. \tag{1}$$

This holds for every $n \geq k$ so that

$$\lim_{n\to\infty} g_n \geq \lim_{n\to\infty} f_{kn} = f_k.$$

But k was an arbitrary positive integer, so it must be the case that

$$\lim_{n\to\infty} g_n \geq f_k \quad \text{for } \underline{\text{every}} \text{ positive integer } k.$$

Hence $\lim\limits_{n\to\infty} g_n \geq \lim\limits_{k\to\infty} f_k$.

Also, for each positive integer n,

$$g_n = f_{1n} \vee \cdots \vee f_{nn} \leq f_1 \vee \cdots \vee f_n = f_n. \tag{2}$$

Thus $\lim\limits_{n\to\infty} g_n \leq \lim\limits_{n\to\infty} f_n$.

We conclude, then, that $\lim\limits_{n\to\infty} g_n = \lim\limits_{n\to\infty} f_n = f$. This shows that f is the limit function of a monotone increasing sequence of functions in L. Thus $f \in U$.

Now in (1) we have, for fixed $k \geq 1$,

$$g_n \geq f_{kn} \quad \text{for every } n \geq k.$$

By property (2') we have

$$I(g_n) \geq I(f_{kn}) \quad \text{for every } n \geq k.$$

Hence $\lim\limits_{n\to\infty} I(g_n) \geq \lim\limits_{n\to\infty} I(f_{kn}) = I(f_k)$.

But this can be done for any $k \geq 1$. Consequently

$$\lim_{n\to\infty} I(g_n) \geq \lim_{k\to\infty} I(f_k).$$

Thus

$$I(f) \geq \lim_{k\to\infty} I(f_k).$$

In (2) we have $g_n \leq f_n$ for each positive integer n. By the extension of property (2') to U we have

$$I(g_n) \leq I(f_n) \text{ for each } n \geq 1.$$

Thus $I(f) = \lim_{n\to\infty} I(g_n) \leq \lim_{n\to\infty} I(f_n)$.

We now have

$$I(f) \leq \lim_{n\to\infty} I(f_n) \leq I(f)$$

which implies the desired result that

$$\lim_{n\to\infty} I(f_n) = I(f).$$

Since $\{f_n\}$ is an increasing sequence of functions in U, $\{I(f_n)\}$ is an increasing sequence by (2'). Hence we write

$$I(f_n)\uparrow I(f).$$

This completes the proof.

We now extend the definition of I to limits of monotone decreasing sequences in L. Note that if $f_n \varepsilon L$ and $f_n \downarrow f$, then $-f_n \varepsilon L$ and $(-f_n) \uparrow (-f)$. With this in mind we define

$$-U = \{f : -f \varepsilon U\}.$$

For $f \varepsilon -U$ we define $I(f) = -I(-f)$. (Note that if f is also in U, this definition agrees with the old one, for $f + (-f) = 0$ and $I(f) + I(-f) = I(0) = 0$ so that $I(f) = -I(-f)$.)

Clearly $-U$ has properties analogous to those of U. That is, $-U$ is closed under monotone decreasing limits, the lattice operations, addition, multiplication by non-negative constants, and I on $-U$ has properties (1) and (2'). Note, also, that if $g \varepsilon -U$, $h \varepsilon U$, and $g \leq h$, then $h-g \varepsilon U$ and

$$I(h) - I(g) = I(h-g) \geq 0.$$

2.4 Definition: A function f mapping X into the extended real line $R^1 \cup \{-\infty, +\infty\}$ is *summable (I-summable)* if and only if for every $\varepsilon > 0$ there exist functions $g \varepsilon -U$ and $h \varepsilon U$ such that $g \leq f \leq h$, $I(g) < \infty$, $I(h) < \infty$, and $I(h) - I(g) < \varepsilon$.

If f is summable, we see that

$$\inf \{I(h) : h \varepsilon U, h \geq f\} = \sup \{I(g) : g \varepsilon -U, g \leq f\}$$

and we define $I(f)$ to be this common value. We will let L^1 (or $L^1(I)$) denote the class of all summable functions. Note that if $f \varepsilon U$ and $I(f) < \infty$, then $f \varepsilon L^1$ and the new definition of $I(f)$ agrees with the old.

2.5 Theorem: The extension of I is an integral on L^1.

Proof: Let $f_1, f_2 \varepsilon L^1$, and let $\varepsilon > 0$ be given. By definition of L^1 there exist functions $g_1, g_2 \varepsilon -U$ and $h_1, h_2 \varepsilon U$ such that $g_i \leq f_i \leq h_i$ and $I(h_i) - I(g_i) < \frac{\varepsilon}{2}$, $i = 1,2$. Let $*$ denote any one of the operations V, Λ, $+$. We have at once that

$$g_1 * g_2 \leq f_1 * f_2 \leq h_1 * h_2. \tag{1}$$

In addition, it is true that

$$(h_1 * h_2) - (g_1 * g_2) \leq (h_1 - g_1) + (h_2 - g_2). \tag{2}$$

Suppose, for instance, that $*$ denotes V. Then inequality (2) is equivalent to the inequality

$$(h_1 V h_2) - (h_1 + h_2) \leq (g_1 V g_2) - (g_1 + g_2). \tag{3}$$

If α and β are any real numbers, it is easily verified that

$$(\alpha+\beta) - (\alpha V \beta) = \alpha \Lambda \beta.$$

Thus (3) is equivalent to the inequality

$$-(h_1 \wedge h_2) \leq -(g_1 \wedge g_2),$$

which is true by (1) with * used as $\wedge$. Thus (2) is true when * denotes V. A similar argument applies if * denotes $\wedge$, and if * denotes +, equality holds in (2). Thus (2) is valid in all cases.

Therefore

$$I(h_1 * h_2) - I(g_1 * g_2) \leq I(h_1 - g_1) + I(h_2 - g_2)$$
$$< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

This shows that $f_1 * f_2 \varepsilon L^1$. That is, $f_1 + f_2$, $f_1 V f_2$, and $f_1 \wedge f_2$ are all in L^1. Since I is additive on U we have

$$|I(f_1+f_2)-I(f_1)-I(f_2)| = |I(f_1+f_2)-I(h_1+h_2)+I(h_1)+I(h_2)-I(f_1)-I(f_2)|$$
$$\leq |I(f_1+f_2)-I(h_1+h_2)| + |I(h_1)-I(f_1)|$$
$$+ |I(h_2)-I(f_2)| < \varepsilon + \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = 2\varepsilon.$$

But $\varepsilon > 0$ was arbitrary. Hence we conclude that $I(f_1+f_2) = I(f_1) + I(f_2)$.

Now let $f \varepsilon L^1$, and let c be any positive scalar. There exist functions g and h such that $g \varepsilon -U$, $h \varepsilon U$, $g \leq f \leq h$, and $I(h) - I(g) < \frac{\varepsilon}{c}$. Then $cg \varepsilon -U$, $ch \varepsilon U$, $cg \leq cf \leq ch$, and $I(ch) - I(cg) = c[I(h)-I(g)] < c\,\frac{\varepsilon}{c} = \varepsilon$. This shows that $cf \varepsilon L^1$ whenever $c > 0$. Clearly, $cf \varepsilon L^1$ if $c = 0$. Note that, for $c < 0$, the roles of cg and ch will be reversed, and a proof similar to that for $c > 0$ establishes that $cf \varepsilon L^1$. We now conclude that I is linear on L^1. That is, property (1) of Definition 2.1 is satisfied on L^1.

If $f \in L^1$ and $f \geq 0$, then any suitable $h \in U$ in the definition of L^1 is such that $0 \leq f \leq h$, and by property (2') on U, we have $I(h) \geq 0$. Hence $I(f) = \inf \{I(h) : h \in U, h \geq f\} \geq 0$. This proves property (2) of Definition 2.1 on L^1. Property (3) of Definition 2.1 is a consequence of the following more general theorem.

2.6 Theorem. (Lebesgue's Monotone Convergence Theorem): If $f_n \in L^1$ $(n=0,1,2,\ldots)$, $f_n \uparrow f$, and $\lim_{n\to\infty} I(f_n) < \infty$, then $f \in L^1$ and $I(f_n) \uparrow I(f)$.

Proof: First we consider the special case for which $f_o \equiv 0$. Put

$$F_n = f_n - f_{n-1} \qquad (n=1,2,\ldots),$$

and note that $F_n \in L^1$, $F_n \geq 0$, and $F_n \to 0$ as $n \to \infty$. Let $\varepsilon > 0$ be given. By definition of a summable function we can choose functions $h_n \in U$ such that $0 \leq F_n \leq h_n$ and $I(h_n) - I(F_n) < \frac{\varepsilon}{2^n}$ $(n=1,2,\ldots)$. Since $f_o \equiv 0$, we see that

$$f_n = \sum_{i=1}^{n} F_i \leq \sum_{i=1}^{n} h_i \tag{1}$$

and

$$\sum_{i=1}^{n} I(h_i) < \sum_{i=1}^{n} [I(F_i) + \frac{\varepsilon}{2^i}] < I(f_n) + \varepsilon. \tag{2}$$

Now put $h = \lim_{n\to\infty} \sum_{i=1}^{n} h_i = \sum_{i=1}^{\infty} h_i$. By Lemma 2.3 it follows that $h \varepsilon U$ and $I(\sum_{i=1}^{n} h_i) \uparrow I(h)$. By the linearity of I we see that $\sum_{i=1}^{n} I(h_i) \uparrow I(h)$. Thus there is an integer $N > 0$ such that

$$I(h) - \sum_{i=1}^{N} I(h_i) < \varepsilon$$

and by (2)

$$I(h) < \sum_{i=1}^{N} I(h_i) + \varepsilon < I(f_N) + 2\varepsilon.$$

Since f_N is a summable function, there is a function $g \varepsilon\text{-}U$ such that $g \leq f_N$ and $I(f_N) - I(g) < \varepsilon$. Taking limits in (1) shows that $f \leq h$. Thus

$$g \leq f_N \leq f \leq h$$

and

$$I(h) - I(g) = [I(h)-I(f_N)] + [I(f_N)-I(g)] < 2\varepsilon + \varepsilon = 3\varepsilon.$$

It follows that $f \varepsilon L^1$ and that $I(f_n) \uparrow I(f)$.

The general case may now be handled by applying the above argument to the sequence $\{f_n - f_o\}$. Since $(f_n - f_o) \uparrow (f - f_o)$, and the initial function in this sequence is identically zero, we have shown that $(f-f_o) \varepsilon L^1$ and $I(f_n - f_o) \uparrow I(f-f_o)$. Since $f_o \varepsilon L^1$ and L^1 is linear, we have

$f \epsilon L^1$ and $I(f_n) \uparrow I(f)$. This completes the proof.

Note that analogous assertions can be made concerning limits of monotone decreasing sequences in L^1. If $f_n \epsilon L^1$ $(n=0,1,\ldots)$ and $f_n \downarrow f$, then $-f_n \epsilon L^1$ $(n=0,1,\ldots)$ and $(-f_n) \uparrow (-f)$. If it is the case that $\lim_{n\to\infty} I(-f_n) < +\infty$ (that is, if $\lim_{n\to\infty} I(f_n) > -\infty$), then by the monotone convergence theorem $-f \epsilon L^1$ and $I(-f_n) \uparrow I(-f)$. By linearity, this asserts that $f \epsilon L^1$ and $I(f_n) \downarrow I(f)$. (In particular, I has property (3) of Definition 2.1 on L^1.)

At this point one might wonder if the hypotheses $\phi_n \epsilon L^1$ and $\phi_n \to \phi$ (not necessarily monotonically) imply that $I(\phi_n) \to I(\phi)$. That this is not the case is shown by considering the following counterexample. For $n = 1,2,\ldots$, we define

$$\phi_n(x) = \begin{cases} n \sin nx, & 0 \leq x \leq \frac{\pi}{n} \\ 0 \quad , & \frac{\pi}{n} < x \leq \pi. \end{cases}$$

Note that $\phi_n(x)$ converges to zero for every $x \epsilon [0,\pi]$. Thus $I(\phi) = I(0) = 0$. (Here, we take I to be the ordinary Riemann integral, $X = [0,1]$, and L as the class of continuous functions on $[0,1]$.) But, for each n,

$$I(\phi_n) = \int_0^{\frac{\pi}{n}} n \sin nx \, dx = [-\cos nx]_0^{\frac{\pi}{n}} = 2.$$

Hence we have

$$\lim_{n\to\infty} I(\phi_n) = 2 \neq I(\phi).$$

The requirement of monotonicity may be replaced, however. (See Lebesgue's dominated convergence theorem in the following section.)

Section 3: The Baire Functions

In Section 2 we extended I to the set of limits of certain types of monotone sequences. We now consider the problem of extending I to a space which contains L and which is closed under the operation of taking monotone limits.

2.7 Definition: A class of real-valued functions is said to be a *monotone class* if it is closed under the operations of taking monotone increasing and monotone decreasing limits of sequences in the class.

2.8 Definition: The *Baire functions* are the members of the smallest monotone class containing the vector lattice L. We denote the class of Baire functions by B.

2.9 Theorem: B is closed under addition, scalar multiplication, and the lattice operations.

Proof: For any $f \in B$, let $M(f)$ be the set of all functions, $g \in B$ such that $f+g$, $f \vee g$, and $f \wedge g$ are in B. For any monotone sequence $\{g_n\}$ of members of $M(f)$, the sequences $\{f+g_n\}$, $\{f \vee g_n\}$, and $\{f \wedge g_n\}$ are monotone sequences in B. If $g_n \to g$, then $f+g$, $f \vee g$, and $f \wedge g$ are in B since B is a monotone class. Thus $g \in M(f)$. This shows that $M(f)$ is a monotone class. Let h be any member of L. If $f \in L$, then $f+h$, $f \vee h$, and $f \wedge h$ are in $L \subset B$. Thus $h \in M(f)$. We conclude, if $f \in L$, that $M(f)$ is a monotone class containing L and hence $B \subset M(f)$. But by definition of $M(f)$,

$M(f) \subset B$. Hence $M(f) = B$ if $f \varepsilon L$. More generally, suppose $f \varepsilon B$. Then for every $h \varepsilon L$, $f \varepsilon M(h)$, since $M(h) = B$. But clearly $f \varepsilon M(h)$ if and only if $h \varepsilon M(f)$. Hence $L \subset M(f)$, and it follows as before that $M(f) = B$. That is, for any f and g in B, $f+g$, $f \vee g$, and $f \wedge g$ are in B.

Let M be the class of all $f \varepsilon B$ for which $cf \varepsilon B$ for every real c. For any monotone sequence $\{f_n\}$ of members of M, $\{f_n\}$ is a monotone sequence in B, as is $\{cf_n\}$ for any real number c. Let $f = \lim_{n \to \infty} f_n$. Since B is a monotone class $f \varepsilon B$ and $cf \varepsilon B$. But c was an arbitrary real number, and hence we conclude that $f \varepsilon M$ and, thus, that M is a monotone class. Now if $f \varepsilon L \subset B$, we have $cf \varepsilon L \subset B$, since L is a linear space. This shows that if $f \varepsilon L$ then $f \varepsilon B$ and $cf \varepsilon B$ for every real c. Thus $f \varepsilon M$, and $L \subset M$. We now have that M is a monotone class including L, from which we conclude that $B \subset M$. But $M \subset B$ by definition of M. Hence $M = B$ and $cf \varepsilon B$ for any $f \varepsilon B$ and any real number c. This completes the proof.

For any class C of functions we let C^+ denote the class of all non-negative functions in C.

2.10 Definition: A function f is said to be *L-bounded* if there exists some $g \varepsilon L^+$ such that $|f| \leq g$. A family F of functions is called *L-monotone* if whenever $\{f_n\}$ is a sequence of L-bounded functions in F, and $f_n \uparrow f$ or $f_n \downarrow f$, then $f \varepsilon F$.

2.11 Lemma: If $f \varepsilon B$, then there exists $g \varepsilon U$ such that $f \leq g$.

Proof: Let T be the class of all $f \varepsilon B$ for which there is some $g \varepsilon U$ such that $f \leq g$. Let $\{f_n\}$ be a monotone sequence of members of T, and let $\{g_n\}$ be a monotone sequence of members of U such that $f_n \leq g_n$, $n = 1,2,\ldots$. (That the sequence $\{g_n\}$ can be chosen monotone is seen by first choosing a sequence $\{h_n\}$ in U such that $f_n \leq h_n$, $n = 1,2,\ldots$. Now put $g_n = h_1 \vee \cdots \vee h_n$ for each n.) Then

$$f = \lim_{n\to\infty} f_n \leq \lim_{n\to\infty} g_n = g,$$

and $g \varepsilon U$ by Lemma 2.3. Thus $f \varepsilon T$. This shows that T is a monotone class. If $h \varepsilon L \subset B$, then $|h| \varepsilon L \subset U$ and $h \leq |h|$. This implies that $h \varepsilon T$, and we conclude that $T \supset L$. That is, T is a monotone class including L. Hence $B \subset T$. But by the definition of T it is true that $T \subset B$. Thus $T = B$, and the proof is complete.

2.12 Lemma: The smallest monotone class containing L^+ is B^+.

Proof: Let $M(L^+)$ denote the smallest monotone class containing L^+. Since B^+ is a monotone class containing L^+, it is clear that

$$M(L^+) \subseteq B^+.$$

Now let M be *any* monotone class including L^+, and define

$$C = \{f : f \varepsilon B,\ f^+ \varepsilon M\}.$$

Suppose $\{f_n\}$ is a monotone sequence in C with limit f. Then each $f_n \varepsilon B$ so that $f \varepsilon B$, and $\{f_n^+\}$ is a monotone sequence in M with limit f^+. Since M is a monotone class, $f^+ \varepsilon M$ and, therefore, $f \varepsilon C$. This shows that C is a monotone class. If $g \varepsilon L$, then $g \varepsilon B$ and $g^+ \varepsilon L^+ \subset M$, so that $g \varepsilon C$. Thus $C \supset L$ and hence $C \supset B$, since $B = M(L)$. We conclude that $B = C$, and thus that

$$\mathcal{B}^+ = \mathcal{C}^+ = \{f : f\varepsilon\mathcal{B}^+,\ f\varepsilon\mathcal{M}\} \subset \mathcal{M}.$$

In particular, $\mathcal{M}(L^+)$ is a monotone class including L^+. Consequently

$$\mathcal{B}^+ \subset \mathcal{M}(L^+).$$

Therefore $\mathcal{B}^+ = \mathcal{M}(L^+)$.

2.13 Theorem: The smallest L-monotone class containing L^+ is $\mathcal{B}^+$.

Proof: Let C be this smallest class, and let g be an arbitrary but fixed member of L^+. Put

$$\mathcal{M} = \{f : f\varepsilon\mathcal{B}^+,\ f\wedge g\varepsilon C\}.$$

If $\{f_n\}$ is a monotone sequence in $\mathcal{M}$, then $\{f_n\wedge g\}$ is a monotone sequence of L-bounded functions in C since for each n, $|f_n\wedge g| = f_n\wedge g \leq g\varepsilon L^+$. Since C is an L-monotone class, $f_n\wedge g \to f\wedge g\varepsilon C$, and hence $f_n \to f\varepsilon\mathcal{M}$. Thus $\mathcal{M}$ is a monotone class. If $f\varepsilon L^+ \subset B^+$, then $f\wedge g\varepsilon L^+ \subset C$, and hence $f\varepsilon\mathcal{M}$. Thus $\mathcal{M}$ is a monotone class including L^+. Hence $\mathcal{M} \supset \mathcal{B}^+$ by Lemma 2.12. Since $\mathcal{M} \subset \mathcal{B}^+$ by definition of $\mathcal{M}$, we conclude that $\mathcal{M} = \mathcal{B}^+$. Thus if $f\varepsilon\mathcal{B}^+$ and $f \leq g$, then $f = f\wedge g\varepsilon C$; that is, C contains every L-bounded function in $\mathcal{B}^+$. Now let f be any member of $\mathcal{B}^+$, and choose $g\varepsilon U$ (by Lemma 2.11) such that $f \leq g$. There exist functions $g_n\varepsilon L^+$ such that $g_n\uparrow g$. Then each $f\wedge g_n\varepsilon C$ by virtue of being an L-bounded member of $\mathcal{B}^+$. Also, $(f\wedge g_n)\uparrow f$. Thus $f\wedge g = f\varepsilon C$, since C is an L-monotone class. We have proved that

$B^+ \subset C$. Since B^+ is itself an L-monotone class containing L^+, and since C is the smallest such, it follows that $B^+ = C$. This completes the proof.

We now replace L^1 by $L^1 \cap B$. That is, from now on a function is not considered summable unless it satisfies Definition 2.4 and is <u>also</u> a Baire function. This restriction to Baire functions is entirely a matter of convenience. It avoids the necessity of various "measure zero" arguments in proofs such as that of the Fubini theorem.

<u>2.14 Theorem</u>: $f \varepsilon L^1$ if and only if $f \varepsilon B$ and $|f| \leq g$ for some $g \varepsilon L^1$.

Proof: If $f \varepsilon L^1$ then $f \varepsilon B$ by our convention. In addition, $|f| \varepsilon L^1$, so that the condition that $|f| \leq g$ for some $g \varepsilon L^1$ is satisfied by taking $g = |f|$. Suppose that $f \varepsilon B$ and $|f| \leq g$ for some $g \varepsilon L^1$. Put

$$M = \{h : h \varepsilon B^+, h \wedge g \varepsilon L^1\}.$$

We see at once that M is a monotone class. For, if $\{h_n\}$ is a monotone sequence in M with limit h, then $\{h_n \wedge g\}$ is a monotone sequence in L^1 with limit $h \wedge g$. Since $\lim_{n \to \infty} I(h_n \wedge g) \leq I(g) < +\infty$, the monotone convergence theorem guarantees that $h \wedge g \varepsilon L^1$. Clearly $h \varepsilon B^+$, since each $h_n \varepsilon B^+$. This shows that $h \varepsilon M$, and hence that M is a monotone class.

Now if $h \varepsilon L^+$ then $h \varepsilon B^+ \cap L^1$. Thus $h \wedge g \varepsilon L^1$ and hence $h \varepsilon M$. We now have $L^+ \subset M$. Consequently M is a monotone class containing L^+ and, therefore, M contains B^+ (the minimal monotone class including L^+). But $M \subset B^+$ by the definition of M. Hence we conclude that $M = B^+$. Thus $f = f \wedge g \varepsilon L^1$, and the proof is complete.

In what follows, the notation "$\underline{\lim}$" denotes "$\liminf\limits_{n\to\infty}$."

<u>2.15 Lemma</u>. (Fatou's Lemma): Let $\{f_n\}$ be a sequence of non-negative functions in L^1. If $\underline{\lim}\, I(f_n) < +\infty$, then $\underline{\lim}\, f_n$ is in L^1 and

$$I(\underline{\lim}\, f_n) \leq \underline{\lim}\, I(f_n).$$

Proof: Let $g_n = f_1 \wedge f_2 \wedge \cdots \wedge f_n$. Then $\{g_n\}$ is a sequence of non-negative functions in L^1 which decreases to $g = \inf\limits_n f_n$. Thus $g_n \downarrow g$ and $I(g_n) \geq 0$. The monotone convergence theorem guarantees that $g \in L^1$. (See the remark following Theorem 2.6.) Put

$$h_k = \inf_{n \geq k} f_n.$$

Then $\{h_k\}$ is a sequence of non-negative functions in L^1 which increases to $\underline{\lim}\, f_n$. Since $h_k \leq f_n$ for each n such that $n \geq k$, we have

$$I(h_k) \leq I(f_n) \quad \text{for each n such that } n \geq k.$$

Thus

$$I(h_k) \leq \inf_{n \geq k} I(f_n),$$

and

$$\lim_{k\to\infty} I(h_k) \leq \lim_{k\to\infty} \inf_{n \geq k} I(f_n) = \underline{\lim}\, I(f_n) < +\infty.$$

Hence by Theorem 2.6, $\lim_{k\to\infty} h_k \varepsilon L^1$--that is, $\underline{\lim}\, f_n \varepsilon L^1$--and

$$I(\underline{\lim}\, f_n) = I(\lim_{k\to\infty} h_k) = \lim_{k\to\infty} I(h_k) \le \underline{\lim}\, I(f_n).$$

This completes the proof.

2.16 Theorem. (Lebesgue's Dominated Convergence Theorem): If $\{f_n\}$ is a sequence of functions in L^1 such that

i) $|f_n| \le g$ $(n=1,2,\ldots)$ for some function $g\varepsilon L^1$, and

ii) $f = \lim_{n\to\infty} f_n$,

then

$$I(f) = \lim_{n\to\infty} I(f_n).$$

Proof: Since $-f_n \le |f_n| \le g$, we have $0 \le g + f_n$ for all positive integers n. Also $|g+f_n| \le 2g\varepsilon L^1$ for all n. Now $\{g+f_n\}$ is a sequence of non-negative functions in L^1, and $0 \le g + f_n \le 2g$. Consequently

$$0 = I(0) \le I(g+f_n) \le I(2g) = 2I(g) < +\infty$$

for each positive integer n. Therefore $\underline{\lim}\, I(g+f_n) < +\infty$. Noting that $\underline{\lim}\,(g+f_n) = \lim(g+f_n) = g + f$, we apply Fatou's lemma to deduce that $g+f \varepsilon L^1$ and

$$I(g+f) \le \underline{\lim}\, I(g+f_n) = I(g) + \underline{\lim}\, I(f_n).$$

Hence

$$I(f) \leq \underline{\lim}\, I(f_n).$$

Since the functions $g-f_n$ are also non-negative members of L^1, we again use Fatou's lemma which asserts that

$$I(g-f) \leq \underline{\lim}\, I(g-f_n) = I(g) + \underline{\lim}\, I(-f_n) = I(g) - \overline{\lim}\, I(f_n)$$

where "$\overline{\lim}$" denotes "$\limsup\limits_{n\to\infty}$." Hence

$$\overline{\lim}\, I(f_n) \leq I(f).$$

We now have

$$I(f) \leq \underline{\lim}\, I(f_n) \leq \overline{\lim}\, I(f_n) \leq I(f).$$

Hence

$$\underline{\lim}\, I(f_n) = \overline{\lim}\, I(f_n) = \lim_{n\to\infty} I(f_n) = I(f).$$

This completes the proof.

We now extend I to any function in B^+ by putting $I(f) = +\infty$ if $f \varepsilon B^+ - L^1$.

2.17 Definition: A function $f \varepsilon B$ is declared *integrable* if either its positive part $f^+ = f \vee 0$, or its negative part $f^- = -(f \wedge 0)$, is summable. For any integrable function $f \varepsilon B$ we define $I(f) = I(f^+) - I(f^-)$.

Note that $+\infty$ and $-\infty$ are possible values of $I(f)$. Also, a function f is summable if and only if both f^+ and f^- are summable. Thus f is summable if and only if f is integrable and $I(|f|) < +\infty$.

2.18 Theorem: If f is any integrable function, then $|I(f)| \leq I(|f|)$.

Proof: If f is integrable but not summable, $I(|f|) = +\infty$ and the result is immediate. If f is summable, then

$$|I(f)| = |I(f^+)-I(f^-)| \leq I(f^+) + I(f^-) = I(|f|)$$

since $I(f^+)$ and $I(f^-)$ are non-negative real numbers.

2.19 Theorem: (a) If f and g are integrable, then $f + g$ is integrable and $I(f+g) = I(f) + I(g)$, provided that $I(f)$ and $I(g)$ are not oppositely infinite.

(b) If f_n is integrable $(n=1,2,3,\ldots)$, $I(f_1) > -\infty$ and $f_n \uparrow f$, then f is integrable and $I(f_n) \uparrow I(f)$.

Proof of (a): If f and g are both summable, the desired result is already guaranteed by Theorem 2.5. Since all other cases can be handled similarly, we shall give a proof for only one of the possible situations. Suppose that $g \varepsilon L^1$ and f is integrable with $I(f) = +\infty$. In this case we must have $I(f^+) = +\infty$ while $I(f^-) < +\infty$. Note that

$$(f+g)^- = -[(f+g)\wedge 0] \le -[(f\wedge 0) + (g\wedge 0)] = f^- + g^- .$$

Therefore

$$I\left((f+g)^-\right) \le I(f^-) + I(g^-) < +\infty,$$

and $f + g$ is integrable.

Now $I(f) + I(g) = +\infty$, and

$$I(f+g) = I\left((f+g)^+\right) - I\left((f+g)^-\right)$$

where $I\left((f+g)^-\right) < +\infty$. If we assume that $I\left((f+g)^+\right) < +\infty$, then $f + g$ is summable as is $(f+g) - g = f$, giving a contradiction. We conclude, then, that $I\left((f+g)^+\right) = +\infty$. Thus $I(f+g) = +\infty$, and it follows that

$$I(f+g) = I(f) + I(g).$$

Part (b) is a simple extension of Theorem 2.6.

Section 4: Equivalence and Measurability

In this section we shall show that I generates a positive measure on X. Also, some results regarding null sets and null functions will be discussed.

2.20 Definition: If $A \subset X$, we define the *characteristic function* of A by

$$\phi_A(x) = \begin{cases} 1, & x\varepsilon A \\ 0, & x\varepsilon A^c. \end{cases}$$

2.21 Definition: A subset A of X is an *integrable set* if and only if $\phi_A \varepsilon \mathcal{B}$.

(Note that $\phi_A \varepsilon \mathcal{B}$ implies that ϕ_A is integrable since $\phi_A^- = 0\varepsilon L^1$.)

2.22 Definition: We define the set function μ by $\mu(A) = I(\phi_A)$ for each integrable set A.

Note that $\mu(\phi) = I(0) = 0$. We shall see presently that μ is, in fact, a positive measure.

2.23 Theorem: If A and B are integrable sets, then so are the sets $A \cup B$, $A \cap B$, and $A-B$.

Proof: The result follows at once by observing that

$$\phi_{A\cup B} = \phi_A \vee \phi_B,\ \phi_{A\cap B} = \phi_A \wedge \phi_B,\ \text{and}\ \phi_{A-B} = \phi_A - \phi_{A\cap B}.$$

Remark: If A and B are integrable sets with $A \subset B$, then $\phi_A \le \phi_B$, and thus $\mu(A) \le \mu(B)$. The set function μ is *monotone*.

2.24 Theorem: If $\{A_n\}$ is a sequence of pairwise-disjoint integrable sets, and if $A = \bigcup_{n=1}^{\infty} A_n$, then A is integrable and

$$\mu(A) = \sum_{n=1}^{\infty} \mu(A_n).$$

Proof: Put $E_1 = A_1$, $E_2 = A_1 \cup A_2, \ldots, E_n = A_1 \cup A_2 \cup \cdots \cup A_n$, ... Note that E_j is integrable, that $E_j \subset E_{j+1}$ for any positive integer j, and that $\bigcup_{j=1}^{\infty} E_j = A$. Thus $\phi_{E_j} \leq \phi_{E_{j+1}}$, and $\phi_{E_j} \uparrow \phi_A$. Now if $\lim_{j\to\infty} I(\phi_{E_j}) < +\infty$, then $I(\phi_{E_j}) < +\infty$ for each j = 1,2,..., and hence each $\phi_{E_j} \varepsilon L^1$. Hence the monotone convergence theorem asserts that $\phi_A \varepsilon L^1$ and $I(\phi_{E_j}) \uparrow I(\phi_A)$. That is, A is integrable (in fact, summable), and

$$\mu(A) = I(\phi_A) = \lim_{j\to\infty} I(\phi_{E_j}) = \lim_{j\to\infty} \sum_{n=1}^{j} I(\phi_{A_n})$$

$$= \sum_{n=1}^{\infty} I(\phi_{A_n}) = \sum_{n=1}^{\infty} \mu(A_n).$$

On the other hand, if $\lim_{j\to\infty} I(\phi_{E_j}) = +\infty$, then to each M > 0 there corresponds an integer N > 0 such that

$$I(\phi_A) \geq \sum_{n=1}^{j} I(\phi_{A_n}) = I(\phi_{E_j}) > M$$

whenever j > N. We conclude, then, that

$$\mu(A) = \sum_{n=1}^{\infty} \mu(A_n) = +\infty.$$

(Clearly, A is integrable since ϕ_A is the limit of a monotone sequence in B, and $\phi_A^- = 0 \varepsilon L^1$.) This completes the proof.

<u>2.25 Corollary:</u> μ is a positive measure.

2.26 Definition: If A is an integrable set and $\mu(A) < \infty$, then A will be called a *summable* set. If the space X is a summable set, the integral I is said to be *bounded*.

We shall now further restrict L by adding a hypothesis used by Stone [7], namely, that

$$f\varepsilon L \Rightarrow (f\wedge 1)\varepsilon L.$$

Note that this hypothesis immediately implies that $f\vee(-1)\varepsilon L$ whenever $f\varepsilon L$. To deduce this, note that if

$$f\varepsilon L,\text{ then } -f\varepsilon L,\ (-f)\wedge 1\varepsilon L,\ -(-f\wedge 1)\varepsilon L$$

and $-(-f\wedge 1) = f\vee(-1)$. Furthermore, these properties are preserved through the extension, so that $f\varepsilon B \Rightarrow f\wedge 1\varepsilon B$.

2.27 Theorem: If $f\varepsilon B$, and $\alpha > 0$, then $A = \{x : f(x) > \alpha\}$ is an integrable set. If $f\varepsilon L^1$, then the set A is summable.

Proof: Let $f\varepsilon B$, and let $\alpha > 0$ be given. For each positive integer n, we define

$$f_n = [n(f-f\wedge\alpha)]\wedge 1.$$

Since $f\varepsilon B$, $\left(\frac{1}{\alpha}\right)f\varepsilon B$ and $\left(\frac{1}{\alpha}\right)f\wedge 1\varepsilon B$. Thus $\alpha[\left(\frac{1}{\alpha}\right)f\wedge 1] = f\wedge\alpha\varepsilon B$. Also $n(f-f\wedge\alpha)\varepsilon B$, since B is a linear space. Hence we conclude that $f_n\varepsilon B$ for each $n = 1,2,\ldots.$ Now if $x\notin A$, then $f(x) \le \alpha$, in which case

$f_n(x) = 0$. But if $x \varepsilon A$, then $f(x) > \alpha$ and $f_n(x) = [n(f(x)-\alpha)] \wedge 1$. Since $f(x) - \alpha > 0$ we let N be the largest integer less than or equal to $1/(f(x)-\alpha)$. Then $f_n(x) = 1$ for each $n \geq N + 1$, and $\lim_{n\to\infty} f_n(x) = 1$. This shows that $f_n \uparrow \phi_A$. Hence $\phi_A \varepsilon \mathcal{B}$, and A is integrable.

Note that, if $x \varepsilon A$, we have $f(x) > \alpha$ and hence $f^+(x) > \alpha$ and $\frac{f^+(x)}{\alpha} > 1 = \phi_A(x)$. If $x \notin A$, $\phi_A(x) = 0$ and $0 \leq \frac{f^+(x)}{\alpha}$ and thus for any x whatsoever

$$0 \leq \phi_A(x) \leq \frac{f^+(x)}{\alpha} .$$

If $f \varepsilon L^1$, then $(f^+/\alpha) \varepsilon L^1$. Thus $\phi_A \varepsilon L^1$ by Theorem 2.14 and, hence, $\mu(A) = I(\phi_A) < +\infty$. This shows that A is summable, and the proof is complete.

2.28 Theorem. If $f \geq 0$, and $A = \{x : f(x) > \alpha\}$ is integrable for every $\alpha > 0$, then $f \varepsilon \mathcal{B}^+$.

Proof: Given $\delta > 1$, let

$$A_m^\delta = \{x : \delta^m < f(x) \leq \delta^{m+1}\}$$

for *every integer* m. Let ϕ_m^δ be the characteristic function of A_m^δ, and let

$$f_\delta = \lim_{N\to\infty} \sum_{m=-N}^{N} \delta^m \phi_m^\delta .$$

Since $A_m^\delta = \{x : f(x) > \delta^m\} - \{x : f(x) > \delta^{m+1}\}$, we see that A_m^δ is integrable for each integer m and, therefore, that $\phi_m^\delta \varepsilon B^+$ for each integer m. Hence the linear combinations

$$f_{\delta,N} = \sum_{m=-N}^{N} \delta^m \phi_m^\delta$$

are in B^+ for each N = 1,2,..., and $\{f_{\delta,N}\}$ is a monotone increasing sequence of functions in B^+. Since B^+ is a monotone class, $f_\delta = \lim_{N\to\infty} f_{\delta,N} \varepsilon B^+$. Since the above can be done for any $\delta > 1$, we choose a suitable sequence $\{\delta_n\}$ such that $\delta_n \downarrow 1$ (say $\delta_n = (n+1)/n$). Then $f_{\delta_n} \uparrow f$ and since each $f_{\delta_n} \varepsilon B^+$, we conclude that $f \varepsilon B^+$. The proof is complete.

<u>2.29 Corollary:</u> If $f\varepsilon B^+$ and $p > 0$, then $f^p \varepsilon B^+$.

Proof: Since $f \geq 0$, then $f^p > \alpha > 0$ if and only if $f > \alpha^{\frac{1}{p}}$. Let

$$A = \{x : f^p(x) > \alpha\} = \{x : f(x) > \alpha^{\frac{1}{p}}\}.$$

Since $f\varepsilon B$ and $\alpha^{\frac{1}{p}} > 0$, A is integrable by Theorem 2.27. Therefore $\{x : f^p(x) > \alpha\}$ is integrable for every $\alpha > 0$. Hence $f^p \varepsilon B^+$ by Theorem 2.28.

<u>2.30 Corollary:</u> If $f\varepsilon B^+$ and $g\varepsilon B^+$, then $fg\varepsilon B^+$.

Proof: Note that $fg = [(f+g)^2-(f-g)^2]/4$. The result follows from this identity.

<u>2.31 Corollary:</u> If $f\varepsilon B^+$, then $I(f) = \int_X f d\mu$, where μ is the measure induced by I and the integral is that of the standard measure approach to integration.

Proof: For $f \varepsilon \mathcal{B}^+$, let f_δ be defined as in the proof of the theorem. Then $f_\delta \leq f \leq \delta f_\delta$, and $I(f_\delta) \leq I(f) \leq \delta I(f_\delta)$. But

$$I(f_\delta) = I(\lim_{N\to\infty} f_{\delta,N}) = \lim_{N\to\infty} I(f_{\delta,N}) = \lim_{N\to\infty} I\left(\sum_{m=-N}^{N} \delta^m \phi_m^\delta\right)$$

$$= \lim_{N\to\infty} \sum_{m=-N}^{N} \delta^m I(\phi_m^\delta) = \lim_{N\to\infty} \sum_{m=-N}^{N} \delta^m \mu(A_m^\delta) = \int_X f_\delta d\mu.$$

Consequently,

$$\int_X f_\delta d\mu \leq I(f) \leq \delta \int_S f_\delta d\mu.$$

We see then that $\int_X f d\mu$ and $I(f)$ are both finite or else both are infinite. If $I(f) = \int_X f d\mu = +\infty$ there is nothing more to prove. Otherwise, we have

$$0 \leq I(f) - \int_X f_\delta d\mu \leq (\delta-1) \int_X f_\delta d\mu \leq (\delta-1)\, I(f).$$

Since δ was an arbitrary number greater than 1, we consider a sequence $\{\delta_n\}$ of numbers such that $\delta_n \downarrow 1$, and note that

$$0 \leq I(f) - \lim_{n\to\infty} \int_X f_{\delta_n} d\mu \leq \lim_{n\to\infty} (\delta_n - 1) I(f) = 0.$$

This shows that $I(f) = \int_X f d\mu$, and completes the proof.

We now direct our attention to some results concerning null functions and null sets so that we shall be able to talk about equivalent functions and L^p-spaces.

2.32 Definition: A function f is said to be *null* if and only if $f \varepsilon B$ and $I(|f|) = 0$. A set A is a *null set* (set of measure zero) if its characteristic function ϕ_A is a null function; that is, A is a null set if A is integrable and its measure is zero.

Some important facts about null functions and null sets are contained in the following theorem.

2.33 Theorem: (a) If g is a null function, and f is any function in B such that $|f| \leq |g|$, then f is a null function.

(b) Any integrable subset of a null set is null.

(c) A countable union of null sets is a null set.

(d) Let E be an integrable set, and let $f = +\infty \cdot \phi_E$ (that is, $f(x) = +\infty$ if $x \varepsilon E$ and $f(x) = 0$ if $x \varepsilon E^c$). Then $f \varepsilon L^1$ if and only if E is a null set, and in this case $I(f) = 0$.

(e) If $g \varepsilon L^1$, then $\{x : |g(x)| = +\infty\}$ is a null set.

Proof of (a): If $0 \leq |f| \leq |g|$, then $I(0) \leq I(|f|) \leq I(|g|)$. But $I(0) = 0$ and, since g is a null function, $I(|g|) = 0$. Hence we have $I(|f|) = 0$. Hence f is a null function.

Proof of (b): Let N be a null set, and suppose E is an integrable subset of N. Then ϕ_N is a null function, $\phi_E \varepsilon B$, and $0 \leq \phi_E \leq \phi_N$. By (a) ϕ_E is a null function. Thus E is a null set.

Proof of (c): Suppose $\{A_n\}$ is a sequence of null sets. Let $A = \bigcup_n A_n$ and $E_n = \bigcup_{i=1}^{n} A_i$. Note that $\phi_{E_n} = \phi_{A_1} \vee \cdots \vee \phi_{A_n}$. It follows that $\phi_{E_n} \varepsilon L^1$. In addition, $0 \leq \phi_{E_n} \leq \phi_{A_1} + \cdots + \phi_{A_n}$, and

$$0 \leq I(\phi_{E_n}) \leq I(\phi_{A_1}) + \cdots + I(\Phi_{A_n}) = 0.$$

Since $\phi_{E_n} \uparrow \phi_A$, we have that $\phi_A \varepsilon L^1$ and $I(\phi_{E_n}) \uparrow I(\phi_A)$, by the monotone convergence theorem. But $I(\phi_{E_n}) = 0$ for each n, and thus we must have $I(\phi_A) = 0$. Therefore A is a null set.

Proof of (d): We define $f_n = n\phi_E$ for each positive integer n. Note that $f_n \uparrow f = +\infty \cdot \phi_E$. If E is a null set, then $I(f_n) = I(n\phi_E) = n \cdot 0 = 0$ for each n and, therefore, $\lim_{n\to\infty} I(f_n) = 0 < +\infty$. Hence by Theorem 2.6 we have that $f\varepsilon L^1$ and $I(f) = \lim_{n\to\infty} I(f_n) = 0$. Conversely, if $f\varepsilon L^1$ then $2f\varepsilon L^1$ and $f = 2f$. Hence $I(f) = 2I(f)$ and $I(|f|) = I(f) = 0$. Since $0 \leq \phi_E \leq f$ where $\phi_E \varepsilon B$ (since E is integrable), and f is a null function, we apply (a) to conclude that ϕ_E is a null function, and hence that E is a null set.

Proof of (e): Given $g\varepsilon L^1$, let $E = \{x : |g(x)| = +\infty\}$. Put $h = +\infty \cdot \phi_E$, and note that $\frac{1}{n}|g| \downarrow h$ and that $I\left(\frac{1}{n}|g|\right) = \frac{1}{n} I(|g|) \downarrow 0$. Then, by the extension of Theorem 2.6 mentioned after that theorem, we have $h\varepsilon L^1$ and $I(h) = 0$. Therefore E is a null set, by (d). This completes the proof.

The following theorem provides a different method of determining whether or not a function in B is null.

<u>2.34 Theorem</u>: A function $f\varepsilon B$ is null if and only if $\{x : f(x) \neq 0\}$ is a null set.

Proof: Let $A = \{x : f(x) \neq 0\}$, and suppose that f is null. Then $n|f|$ is null for each positive integer n since $I(n|f|) = nI(|f|) = 0$. Also $\big|n|f|\wedge 1\big| = n|f|\wedge 1 \leq n|f|$. Hence $n|f|\wedge 1$ is a null function for each positive integer n. If $x\varepsilon A$, then $|f(x)| \neq 0$, and thus there exists a positive integer N such that $n|f(x)| \geq 1$ whenever $n \geq N$. That is, if $x \in A$, then $\lim_{n\to\infty} (n|f|\wedge 1)(x) = 1$. Clearly, if $x \notin A$, then $f(x) = 0$, $(n|f|\wedge 1)(x) = 0$ for all n, and hence $\lim_{n\to\infty} (n|f|\wedge 1)(x) = 0$. This shows that $\lim_{n\to\infty} (n|f|\wedge 1) = \phi_A$. Thus ϕ_A is a Baire function and

$$I(|\phi_A|) = I(\phi_A) = \lim_{n\to\infty} I(n|f|\wedge 1) = \lim_{n\to\infty} 0 = 0.$$

Hence ϕ_A is a null function, and thus A is a null set.

Now suppose A is a null set. Then $+\infty \cdot \phi_A$ is a null function by Theorem 2.33(d). But $0 \leq |f| \leq +\infty \cdot \phi_A$ and hence f is also a null function. This completes the proof.

2.35 Definition: Two functions f and g are said to be *equivalent* if and only if f - g is a null function.

This equivalence relation partitions L^1 into equivalence classes. The following theorem shows that I is constant on the functions of any such equivalence class.

2.36 Theorem: If f and g are equivalent members of L^1, then $I(f) = I(g)$.

Proof: Since f and g are equivalent members of L^1, we apply Theorem 2.18 and the definition of equivalent functions to write

$$0 \leq |I(f)-I(g)| = |I(f-g)| \leq I(|f-g|) = 0.$$

Hence $I(f) = I(g)$.

We now consider a problem that has been intentionally overlooked until now. Since summable functions may assume the values $+\infty$ and $-\infty$, we see that for $f, g \varepsilon L^1$, $f+g$ is not defined on the set where f and g are oppositely infinite. In light of Theorem 2.33(e), however, we see that this "problem" is no problem at all since it only occurs on a set of measure zero. Two standard ways of handling this situation are as follows:

1) Restrict the convergence theorems to *almost everywhere convergence* (pointwise convergence except on a null set), and simply do not worry about the function on the null set in question since it will have no effect whatsoever on the integral.

2) If a function f assumes infinite values on a null set N, agree to replace f by another member of the equivalence class containing f--say, the function $\hat{f}$ that is equal to f on N^c and is zero on N.

In preparing to define measurable sets and functions, we note that X may not itself be an integrable set. Let us assume, then, that $X = \bigcup_{\alpha} X_\alpha$, where $\{X_\alpha\}$ is a disjoint (possibly uncountable) family of integrable sets with the property that every integrable set is included in an at most countable union of the sets X_α.

2.37 Definition: (i) A *set E is measurable* if and only if the sets $E \bigcap X_\alpha$ are all integrable.

(ii) A *function is measurable* if and only if its restriction to X_α is integrable for each α.

CHAPTER III

L^p-SPACES AND THE RADON-NIKODYM THEOREM

Section 1: L^p-Spaces

For the sake of completeness we shall outline here some basic facts about L^p-spaces. The proofs are standard and may be found, for example, in Rudin [5]. For simplicity we restrict attention to the real L^p-spaces.

3.1 Definition: For $1 \leq p < \infty$, let

$$L^p = \{f : f \varepsilon \mathcal{B},\ |f|^p \varepsilon L^1\}.$$

3.2 Minkowski's Inequality: If $f,g\ \varepsilon L^p$ (with p such that $1 \leq p < \infty$), then $f+g\ \varepsilon L^p$ and

$$\left(I(|f+g|^p)\right)^{\frac{1}{p}} \leq \left(I(|f|^p)\right)^{\frac{1}{p}} + \left(I(|g|^p)\right)^{\frac{1}{p}} .$$

The function $\|\cdot\|_p$ defined on L^p by

$$\|f\|_p = \left(I(|f|^p)\right)^{\frac{1}{p}}$$

is a *seminorm* on L^p, since $\|0\|_p = 0$, $\|\alpha f\|_p = |\alpha|\ \|f\|_p$ for any scalar α, and $\|f+g\|_p \leq \|f\|_p + \|g\|_p$ whenever $f,g \varepsilon L^p$ (Minkowski's inequality). Since $\|f\|_p = 0$ implies only that f is a null function and not necessarily the zero function, $\|\cdot\|_p$ is a seminorm, not a norm. Equipped with this seminorm, L^p is a seminormed linear space over the real field. In the study of L^p-spaces, it is customary to introduce the quotient space L^p/N, where N is the subspace of null functions in L^p. The elements of L^p/N are the equivalence classes in L^p induced by the equivalence relation ~ defined by $f \sim g$ if and only if $f-g$ is a null function. If we let $\hat{f}$ denote the element of L^p/N which contains f, then the mapping $\hat{f} \to \|f\|_p$ is well defined, and provides a *norm* on the quotient space, which is then a normed linear space. Following the usual custom, it is convenient to use L^p to denote this normed linear space. This involves an abuse of notation which, however, seldom causes any real difficulty. It is necessary to bear in mind that the elements of L^p are not functions, but equivalence classes of functions.

<u>3.3 Theorem</u>: The normed linear space L^p is *complete* in the sense that if $\{f_n\}$ is a sequence in L^p, such that $\lim_{m,n\to\infty} \|f_m - f_n\|_p = 0$, then there exists $f \in L^p$ such that $\lim_{n\to\infty}\|f_n - f\|_p = 0$. Thus L^p is a complete normed linear space, or a *Banach space*.

<u>3.4 Theorem</u>: L is a dense subset of L^p, if $1 \leq p < \infty$.

<u>3.5 Definition</u>: If the product $fg \in L^1$, we define the *scalar product* (inner product) of f and g by

$$(f,g) = I(fg).$$

<u>3.6 Schwarz Inequality</u>: If $f \in L^2$ and $g \in L^2$, then $fg \in L^1$ and

$$|(f,g)| \leq \|f\|_2\|g\|_2.$$

The complete normed linear space L^2 is of special importance since there exists an inner product which induces the norm in the sense that

$$(f,f) = \|f\|_2^2.$$

The real Banach space L^2 is an important example of a Hilbert space. Hilbert spaces are discussed in Rudin [5], Chapter 4, from an axiomatic point of view, and the space real L^2 is just one very important example of a real Hilbert space. The following fundamental result on Hilbert spaces is important in the proof to be used for the Radon-Nikodym theorem.

3.7 Theorem (Riesz-Fréchet representation theorem in Hilbert spaces): Let H be a Hilbert space, and let $F : H \to R^1$ be a linear functional on H bounded in the sense that

$$\sup \{|F(x)| : \|x\| \leq 1\} < \infty.$$

Then there exists a unique element $y \varepsilon H$ such that

$$F(x) = (x,y) \quad \text{for all } x \varepsilon H.$$

This result, which holds also for complex-valued bounded linear functions on H, is proved in Rudin [5], page 80. We shall use it only for the special case $H = L^2$.

The Banach space L^∞ is also frequently used in analysis. A measurable function f is called *essentially bounded* if and only if there exists a real number k such that $|f(x)| \leq k$ almost everywhere [I]. The symbol L^∞ denotes the collection of all essentially bounded functions, and the mapping $f \to \|f\|_\infty$ defined by

$$\|f\|_\infty = \inf \{k : |f(x)| \leq k \text{ almost everywhere [I]}\}$$

is a seminorm on L^∞. The quotient space L^∞/N can be proved to be a Banach space, also customarily designated by L^∞.

Section 2: The Radon-Nikodym Theorem

We now consider our original vector lattice L as a normed linear space under the uniform norm $\|f\| = \|f\|_\infty$. Since the L^p-spaces are very much dependent upon the integral, we shall write $L^p(I)$, $L^p(J)$, and so forth, when more than one integral is being used.

3.8 Definition: A linear functional F is *bounded* if and only if

$$\sup \{|F(x)| : \|x\| \leq 1\} < \infty.$$

The norm of a bounded linear functional F is defined by

$$\|F\| = \sup \{|F(x)| : \|x\| \leq 1\}.$$

3.9 Theorem: Every bounded linear functional on L may be expressed as the difference of two bounded positive linear functionals.

Proof: Let F be the given bounded linear functional. For each $f \varepsilon L^+$, we define

$$F^+(f) = \sup \{F(g) : 0 \leq g \leq f\}.$$

Note that $F^+(f) \geq 0$ since $g \equiv 0$ is one permissible choice of g, and $F^+(f) = \sup \{F(g) : 0 \leq g \leq f\} \geq F(0) = 0$. In addition,

$$|F^+(f)| = F^+(f) \leq \sup \{\|F\| \cdot g : 0 \leq g \leq f\} = \|F\| \cdot f \leq \|F\| \cdot \|f\|,$$

where the last inequality holds almost everywhere [I]. Clearly $F^+(cf) = cF^+(f)$ if $c > 0$. Consider now a pair of functions $f_1, f_2 \varepsilon L^+$. If g_1 and g_2 are such that $0 \leq g_1 \leq f_1$ and $0 \leq g_2 \leq f_2$, then $0 \leq g_1 + g_2 \leq f_1 + f_2$, and hence

$$F^+(f_1+f_2) \geq F^+(g_1+g_2) = \sup \{F(h) : 0 \leq h \leq g_1 + g_2\}$$

$$\geq F(g_1+g_2) = F(g_1) + F(g_2).$$

Since the above inequalities hold for any g_1 and g_2 such that $0 \leq g_1 \leq f_1$ and $0 \leq g_2 \leq f_2$, we conclude that

$$F^+(f_1+f_2) \geq \sup\{F(g_1) : 0 \leq g_1 \leq f_1\} + \sup\{F(g_2) : 0 \leq g_2 \leq f_2\}$$

$$= F^+(f_1) + F^+(f_2).$$

On the other hand, if $0 \leq g \leq f_1 + f_2$, then $0 \leq f_1 \wedge g \leq f_1$ and $0 \leq g - f_1 \wedge g \leq f_2$, and consequently

$$F^+(f_1+f_2) = \sup\{F(g) : 0 \leq g \leq f_1 + f_2\}$$

$$= \sup\{F(g+f_1\wedge g-f_1\wedge g) : 0 \leq g \leq f_1 + f_2\}$$

$$\leq \sup\{F(f_1\wedge g) : 0 \leq g \leq f_1 + f_2\}$$

$$+ \sup\{F(g-f_1\wedge g) : 0 \leq g \leq f_1 + f_2\}$$

$$\leq \sup\{F(h) : 0 \leq h \leq f_1\} + \sup\{F(h) : 0 \leq h \leq f_2\}$$

$$= F^+(f_1) + F^+(f_2).$$

We now have shown that

$$F^+(f_1) + F^+(f_2) \leq F^+(f_1+f_2) \leq F^+(f_1) + F^+(f_2),$$

which asserts that

$$F^+(f_1+f_2) = F^+(f_1) + F^+(f_2).$$

Thus F^+ is additive on non-negative functions. We now extend F^+ to any $f\varepsilon L$ by defining

$$F^+(f) = F^+(f^+) - F^+(f^-).$$

The functional F^+ is bounded, since

$$|F^+(f)| \leq F^+(f^+) + F^+(f^-) = F^+(|f|) \leq \|F\| \cdot \|f\|$$

almost everywhere [I].

Now let $F^-(f) = F^+(f) - F(f)$. Since $F^+(f) \geq F(f)$ whenever $f \geq 0$, we see that $F^-(f) \geq 0$ whenever $f \geq 0$. Also, F^- is a bounded functional, since both F and F^+ are bounded. Hence we have that F^+ and F^- are bounded positive linear functionals and

$$F = F^+ - F^-.$$

This completes the proof.

3.10 Definition: An integral J is *absolutely continuous* with respect to an integral I (denoted J << I) if and only if every I-null set is also J-null.

3.11 Theorem (Radon-Nikodym Theorem): If the bounded integral J is absolutely continuous with respect to the bounded integral I, then there exists an I-unique, I-summable function f_0 such that ff_0 is I-summable and

$$J(f) = I(ff_0) \quad \text{for every } f \in L^1(J).$$

(To say that f_0 is I-unique means that f_0, f_0^* are I-equivalent if f_0^* is another such function.)

Proof: Consider the bounded integral $K = I + J$, and the real Hilbert space $L^2(K)$. If $f \in L^2(K)$, then by the Schwarz inequality

$$K(|f|) = K(|f|\cdot 1) \leq \|f\|_2 \|1\|_2 < \infty,$$

and hence $f \in L^1(K)$. Note that

$$|J(f)| \leq J(|f|) \leq K(|f|) \leq \|f\|_2 \|1\|_2 < \infty$$

whenever $f \in L^2(K)$. Thus J is a bounded linear functional on $L^2(K)$. By Theorem 3.7 there exists a unique $g \in L^2(K)$ such that

$$J(f) = (f,g) = K(fg) \quad \text{for each } f \in L^2(K). \tag{1}$$

Clearly g is non-negative almost everywhere [K], since J and K are positive linear functionals. By consecutive applications of (1) we see that

$$\begin{aligned} J(f) = K(fg) &= I(fg) + J(fg) \\ &= I(fg) + K(fg^2) \\ &= I(fg) + I(fg^2) + J(fg^2) \\ &= I(fg) + I(fg^2) + K(fg^3) \\ &\cdots\cdots\cdots\cdots \end{aligned}$$

$$= \sum_{i=1}^{n} \left(I(fg^i) \right) + J(fg^n).$$

Now let $E = \{x : g(x) \geq 1\}$. Then, with $f = \phi_E$, it follows that

$$J(\phi_E) = \sum_{i=1}^{n} \left(I(\phi_E g^i) \right) + J(\phi_E g^n). \tag{2}$$

Since $\phi_E \leq \phi_E g^n$, $J(\phi_E) \leq J(\phi_E g^n)$. By this fact, and (2), it follows that

$$\sum_{i=1}^{n} I(\phi_E g^i) = 0.$$

In particular,

$$0 \leq I(\phi_E) \leq I(\phi_E g) = 0,$$

and consequently E is an I-null set. Since J is absolutely continuous with respect to I, E is also a J-null set. Therefore $0 \leq g \leq 1$ almost everywhere with respect to J. Thus $fg^n \downarrow 0$ almost everywhere [J] whenever $f \geq 0$. Since $f \varepsilon L^1(J)$, the monotone convergence theorem asserts that

$$J(fg^n) \downarrow 0. \tag{3}$$

Since I is linear, the expansion

$$J(f) = \sum_{i=1}^{n} \left(I(fg^i) \right) + J(fg^n)$$

may also be written in the form

$$J(f) = I(f \sum_{i=1}^{n} g^i) + J(fg^n)$$

for each positive integer n. Therefore, by the monotone convergence theorem,

$$J(f) = \lim_{n\to\infty} [I(f \sum_{i=1}^{n} g^i) + J(fg^n)] = I(ff_0)$$

where $f_0 = \sum_{i=1}^{\infty} g^i$. Taking $f \equiv 1$, we see that $f_0 \varepsilon L^1(I)$. Also

$$f_0 = \sum_{i=1}^{\infty} g^i = \frac{g}{1 - g} ,$$

and hence

$$g = \frac{f_0}{1 + f_0} .$$

Since g is unique modulo K-null functions, we see that f_0 is unique [I]. Since $J(f) = I(ff_0)$ for each $f \varepsilon L^2(K)$, then, in particular, $J(f) = I(ff_0)$ for each $f \varepsilon L$. Hence the integrals are also identical on $L^1(J)$, and the proof is complete.

CHAPTER IV

EXISTENCE OF LEBESGUE MEASURE IN R^k

In this chapter we shall show that if the original domain space X is taken to be k-dimensional Euclidean space R^k, and if we choose the initial vector lattice L and the Daniell integral I in the appropriate way, then the Daniell method leads to the classical Legesgue integral, and to Lebesgue measure on R^k. We begin by developing an appropriate vector lattice, and defining an integral on the vector lattice.

4.1 Definition: For any function $f : R^k \to R^1$ let $N(f) = \{x : x \in R^k, f(x) \neq 0\}$. The *support of f* is defined to be $\overline{N(f)}$, the closure of $N(f)$. Whenever $\overline{N(f)}$ is a compact set, we say that f has *compact support*.

We shall agree to let $C_c(X)$ denote the class of all continuous real-valued functions on X having compact supports. We define $I : C_c(R^k) \to R^1$ as follows: for any $f \in C_c(R^k)$, let

$$I(f) = \int_E f(x)dx,$$

where E is any compact interval containing $N(f)$ and the integral is the Riemann integral. It is easy to see that $C_c(R^k)$ is a vector lattice. Also, Stone's axiom (page 28) is satisfied here since, if $f \in C_c(R^k)$ and $g = 1 \wedge f$, then g is continuous and $N(g) = N(f)$, and hence $g \in C_c(R^k)$. We now check to see whether we have, in fact, defined a Daniell integral on the vector lattice $C_c(R^k)$. Clearly $I(f) = \int_E f(x)dx$ is independent of the compact interval E containing N(f), since, for instance, if E and F are compact intervals containing N(f), then f(x) = 0 for each $x\varepsilon(E\cap F)^c$, and

$$\int_E f(x)dx = \int_{E\cap F} f(x)dx + \int_{E-F} f(x)dx = \int_{E\cap F} f(x)dx$$

$$= \int_{F\cap E} f(x)dx + \int_{F-E} f(x)dx = \int_F f(x)dx.$$

If $f,g\varepsilon C_c(R^k)$, $\alpha\varepsilon R^1$, $\beta\varepsilon R^1$, and E is a compact interval containing $N(\alpha f)$ and $N(\beta g)$, then

$$I(\alpha f+\beta g) = \int_E (\alpha f+\beta g)dx = \alpha \int_E f(x)dx + \beta \int_E g(x)dx$$

$$= \alpha I(f) + \beta I(g).$$

Thus the linearity of I is ensured by virtue of the linearity properties of the Riemann integral. The fact that I is a positive linear functional is also an immediate consequence of the properties of the Riemann integral. Suppose that $f_n\varepsilon C_c(R^k)$ for each positive integer n, and suppose that $f_n \downarrow 0$. Let E_1 be a compact interval containing $N(f_1)$. Since $f_1 \geq f_2 \geq \cdots \geq f_n \geq \cdots \geq 0$, it is clear that E_1 also contains $N(f_n)$ for each positive integer n. Now on E_1^c each $f_n = 0$. On E_1, $\{f_n\}$ is a monotone sequence of continuous functions mapping the compact set E_1 into R^1. By Dini's theorem (see, for example, Rudin [4], p. 136),

$f_n \to 0$ uniformly on E_1. Hence

$$\lim_{n\to\infty} I(f_n) = \lim_{n\to\infty} \int_{E_1} f_n(x)dx = \int_{E_1} \lim_{n\to\infty} f_n(x)dx = \int_{E_1} 0dx = 0.$$

This shows that I is a Daniell integral.

By the theory developed in Chapter II, I induces a class L^1 of summable functions, a class F of measurable functions, a class S of measurable subsets of R^k, and a positive measure μ on S. (In regard to the discussion of measurability, we shall soon observe that in this particular application the space (R^k, S, μ) is σ-finite. That is, R^k can be written as a countable union of summable [μ] sets X_n.) The measure space (R^k, S, μ) is a complete measure space for if $N \varepsilon S$ with $\mu(N) = 0$, and if $A \subset N$, then $\phi_A \leq \phi_N$, and thus $0 \leq I(\phi_A) \leq I(\phi_N) = \mu(N) = 0$, and $I(\phi_A) = 0$. Therefore, $A \varepsilon S$ and $\mu(A) = 0$.

We shall let M denote the class of sets in R^k which are measurable in the sense of Lebesgue, and we shall denote the Lebesgue measure function by m. A function which is summable in the Lebesgue sense is called L-summable. Sets or functions which are measurable in the Lebesgue sense will be referred to as L-measurable. For facts about the basic Lebesgue measure space (R^k, M, m) reference may be made to Rudin [5], Chapter 2. In the following theorem, R^* denotes the extended real line, and I denotes the extended Daniell integral for the particular example with which we are working.

<u>4.2 Theorem</u>: (a) A function $f : R^k \to R^*$ is I-summable if and only if it is L-summable. If f is I-summable (or L-summable) then

$$I(f) = \int_{R^k} f \, dm.$$

(b) The classes S and M are the same, and the measures μ and m are the same.

(c) A function f is I-measurable if and only if it is L-measurable.

Since the proof of this theorem is rather lengthy, we consider the proof in a number of steps.

<u>Step 1</u>: If E is a finite open interval (k-cell) in R^k, then $E \epsilon S$ and $\mu(E) = m(E)$.

Proof: To prove this assertion, we shall exhibit a sequence $\{f_n\}$ in $C_c(R^k)$ such that $f_n \uparrow \phi_E$. Then the monotone convergence theorem can be applied to obtain the desired result. We construct $\{f_n\}$ as follows. Let the interval E be the set of all points $(x_1,\ldots,x_k) \epsilon R^k$ such that $\alpha_i < x_i < \beta_i$, $i = 1,\ldots,k$. Let e be the length of the shortest edge of E; that is,

$$e = \min \{\beta_i - \alpha_i : i = 1,2,\ldots,k\}.$$

For any $x \epsilon R^k$, let

$$d(x,E^c) = \inf_{y \epsilon E^c} \{|x-y|\}$$

denote the distance from x to E^c. Note that $d(x,E^c) = 0$ if $x \varepsilon E^c$. Also, for each $x \varepsilon R^k$, $0 \leq d(x,E^c) \leq \frac{e}{2}$. For each positive integer n, let

$$f_n(x) = 1 \wedge \frac{2n}{e} d(x,E^c) \quad (x \varepsilon R^k).$$

If $k = 1$, E is an open interval on the real line, and e is the length of the interval. The graph of a typical f_n for this case is shown below.

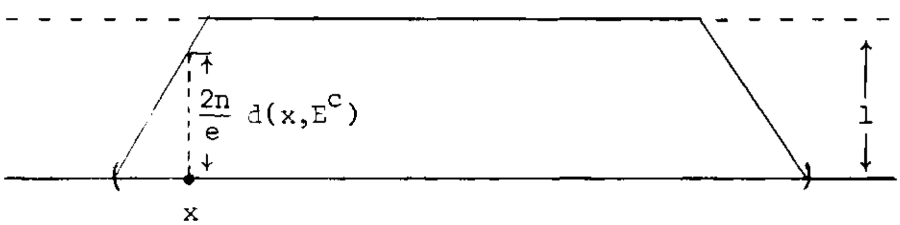

Figure 1. Graph of f_n for the case $k = 1$

Clearly $f_n \varepsilon L^1$ $\{$in particular, $f_n \varepsilon C_c(R^k)\}$ for each positive integer n, $f_n \uparrow \phi_E$, and $\lim_{n\to\infty} I(f_n) < \infty$. By the monotone convergence theorem, $\phi_E \varepsilon L^1$ and $\lim_{n\to\infty} I(f_n) = I(\phi_E) = \mu(E)$. This shows that $E \varepsilon S$. Also, since each f_n is continuous, the Lebesgue integral $\int_{R^k} f_n dm$ agrees with the Riemann integral $I(f_n)$, and the monotone convergence theorem as in the measure approach to integration (see Rudin [5]) implies that

$$m(E) = \int_{R^k} \phi_E dm = \lim_{n\to\infty} \int_{R^k} f_n dm = \lim_{n\to\infty} I(f_n) = \mu(E).$$

We have shown that $E \varepsilon S$ and $\mu(E) = m(E)$ for each finite open interval E in R^k. This completes Step 1.

As a result of Step 1, we see that if A is any bounded set in $\mathcal{S}$, then $\mu(A) < \infty$. (There is some finite open interval E containing A, and hence $\mu(A) \leq \mu(E) < \infty$.) Since R^k can be written as a countable union of bounded sets, it clearly follows that the measure space $(R^k, \mathcal{S}, \mu)$ is σ-finite.

Step 2: If f is I-summable, then f is L-summable and $I(f) = \int_{R^k} f\, dm$.

Proof: Suppose, first, that $h \in U \cap L^1$, where U is the closure of $C_c(R^k)$ under the formation of limits of monotone increasing sequences in $C_c(R^k)$, and L^1 is the final class of summable functions relative to I (or μ). Then there exists a sequence $\{h_n\}$ in $C_c(R^k)$ such that $h_n \uparrow h$ and $I(h_n) \to I(h) < \infty$. Since each h_n is continuous, the Lebesgue integral of h_n is equal to the Riemann integral of h_n, and we infer by the monotone convergence theorem for Lebesgue integrals that h is L-summable and $\int_{R^k} h\, dm = I(h)$.

Now suppose that $f \in L^1$ (that is, that f is I-summable). From the definition of L^1, there is a sequence $\{h_n\}$ of functions in $U \cap L^1$ and an $h \in L^1$ such that $h_n \downarrow h$, $h \geq f$, $I(h_n) \downarrow I(h)$, and $I(h) = I(f)$. By the preceding paragraph, it follows that h is L-summable and $I(h) = \int_{R^k} h\, dm$. If we can show that $f(x) = h(x)$ almost everywhere [m], then we can conclude that f is L-summable and $I(f) = \int_{R^k} f\, dm$. Since $f \leq h$ and $I(f) = I(h)$, we see that

$$0 \leq I(|h-f|) = I(h-f) = I(h) - I(f) = 0.$$

This shows that $f(x) = h(x)$ almost everywhere [μ]. Note that $m \ll \mu$ (m is absolutely continuous with respect to μ) since, if $E \in \mathcal{S}$ with $\mu(E) = 0$, then $\phi_E \in L^1$ and by the first part of this paragraph there is a function h which is L-summable and such that $\phi_E \leq h$ and $\int_{R^k} h\, dm = I(\phi_E) = \mu(E) = 0$. Thus $E \subset \{x : h(x) \neq 0\} = N(h)$. But $m(N(h)) = 0$. Since the measure space $(R^k, \mathcal{M}, m)$ is complete, we conclude that $E \in \mathcal{M}$ and $m(E) = 0$. Now since $m \ll \mu$ and $\mu(\{x : f(x) \neq h(x)\}) = 0$, then

$$m(\{x : f(x) \neq h(x)\}) = 0.$$

This shows that $f(x) = h(x)$ almost everywhere [m], and thus that f is L-summable and $I(f) = \int_{R^k} f\, dm$ for any $f \in L^1$. This concludes the proof of Step 2.

<u>Step 3</u>: $\mathcal{S} \subset \mathcal{M}$ and $\mu(E) = m(E)$ for each $E \in \mathcal{S}$.

Proof: Let $E \in \mathcal{S}$. If $\mu(E) < +\infty$, the result follows from Step 2, by taking $f = \phi_E$. Suppose that $\mu(E) = +\infty$. Since the measure space $(R^k, \mathcal{S}, \mu)$ is σ-finite, we can write $R^k = \bigcup_{n=1}^{\infty} X_n$, where $X_i \cap X_j = \phi$ whenever $i \neq j$, and $\mu(X_n) < +\infty$ for each $n = 1, 2, \ldots$. Then $E = \bigcup_{n=1}^{\infty} (E \cap X_n)$ and $\mu(E \cap X_n) < +\infty$ for each n. If $F_N = \bigcup_{n=1}^{N} (E \cap X_n)$, then ϕ_{F_N} is I-summable for each N. Hence Step 2 implies that ϕ_{F_N} is L-summable and $I(\phi_{F_N}) = \int_{R^k} \phi_{F_N} dm$; that is, $\mu(F_N) = m(F_N)$. Since $F_1 \subset F_2 \subset \cdots$ and $\lim_{N\to\infty} F_N = E$, then $E \in \mathcal{M}$ and

$$\mu(E) = \lim_{N\to\infty} \mu(F_N) = \lim_{N\to\infty} m(F_N) = m(E)$$

by a standard continuity theorem for positive measures. (See, for example, Rudin [5], p. 16.)

Step 4: If $E \epsilon M$ and $m(E) < +\infty$, then $E \epsilon S$ and $\mu(E) = m(E)$.

Proof: We shall demonstrate this in several parts. Note that we already have the result when E is a finite open interval in R^k (or, synonymously, an open k-cell).

Part 1: Suppose $E = E_1 \cup E_2$, where E_1 and E_2 are finite open k-cells. If $E_1 \cap E_2 = \Phi$, then $\phi_E = \phi_{E_1} + \phi_{E_2}$. Otherwise $E_1 \cap E_2$ is another open k-cell and

$$\phi_E = \phi_{E_1} + \phi_{E_2} - \phi_{E_1 \cap E_2}.$$

Hence ϕ_E is a linear combination of members of L^1, and hence $\phi_E \epsilon L^1$, and

$$\begin{aligned} \mu(E) = I(\phi_E) &= I(\phi_{E_1}) + I(\phi_{E_2}) - I(\phi_{E_1 \cap E_2}) \\ &= \mu(E_1) + \mu(E_2) - \mu(E_1 \cap E_2) \\ &= m(E_1) + m(E_2) - m(E_1 \cap E_2) = m(E). \end{aligned}$$

Part 2: Using induction, we extend the result of Part 1 to any finite union of open k-cells. As the induction hypothesis, we assume that if $O = O_1 \cup O_2 \cup \cdots \cup O_n$ where each O_i is an open k-cell, then $O \epsilon S$ and $\mu(O) = m(O)$. Let $E = E_1 \cup E_2 \cup \cdots \cup E_n \cup E_{n+1}$, where each E_i is an open k-cell. Write $E = F \cup E_{n+1}$, where $F = E_1 \cup E_2 \cup \cdots \cup E_n$. Then $\mu(F) = m(F)$ by hypothesis, and

$$\phi_E = \phi_F + \phi_{E_{n+1}} - \phi_{F \cap E_{n+1}}.$$

Note that

$$F \cap E_{n+1} = (E_1 \cap E_{n+1}) \cup (E_2 \cap E_{n+1}) \cup \cdots \cup (E_n \cap E_{n+1}),$$

a union of n open k-cells to which our induction hypothesis applies. Hence $\phi_E \varepsilon L^1$, since ϕ_E is a linear combination of members of L^1. Moreover,

$$\mu(E) = I(\phi_E) = I(\phi_F) + I(\phi_{E_{n+1}}) - I(\phi_{F \cap E_{n+1}})$$

$$= \mu(F) + \mu(E_{n+1}) - \mu(F \cap E_{n+1})$$

$$= m(F) + m(E_{n+1}) - m(F \cap E_{n+1}) = m(F \cup E_{n+1}) = m(E).$$

This shows that if E is any finite union of open k-cells, then $E\varepsilon S$ and $\mu(E) = m(E)$.

Part 3: Next, suppose that $E = \bigcup_{i=1}^{\infty} E_i$, a countable union of open k-cells, where $m(E) < +\infty$. Put $F_n = \bigcup_{i=1}^{n} E_i$. By Part 2, each $F_n \varepsilon S$ and $\mu(F_n) = m(F_n)$. Since $\phi_{F_n} \uparrow \phi_E$ where each $\phi_{F_n} \varepsilon L^1$ and

$$\lim_{n\to\infty} I(\phi_{F_n}) = \lim_{n\to\infty} \mu(F_n) = \lim_{n\to\infty} m(F_n) = m(E) < +\infty,$$

we may apply the monotone convergence theorem to conclude that $\phi_E \varepsilon L^1$ and $I(\phi_E) = \lim_{n\to\infty} I(\phi_{F_n})$. That is, $E\varepsilon S$ and $\mu(E) = m(E)$.

Part 4: Suppose that $E \varepsilon \mathcal{M}$ and $m(E) < +\infty$. For each positive integer n, let $\{E_{ni}\}_{i=1}^{\infty}$ be a countable collection of open k-cells such that $E \subset \bigcup_{i=1}^{\infty} E_{ni}$, and let these open coverings be such that if $F_n = \bigcup_{i=1}^{\infty} E_{ni}$, then $F_n \supset F_{n+1}$, $m(F_1) < +\infty$, and $\lim_{n\to\infty} m(F_n) = m(E)$. (That such open coverings exist is a well-known property of Lebesgue measure spaces.) Put $F = \bigcap_{n=1}^{\infty} F_n = \bigcap_n (\bigcup_i E_{ni})$. By Part 3, each $F_n \varepsilon \mathcal{S}$ and $\mu(F_n) = m(F_n)$. Since $m(F_1) < +\infty$ and $F_n \supset F_{n+1}$, a continuity theorem for positive measures asserts that

$$\lim_{n\to\infty} m(F_n) = m(F).$$

Also, $\mu(F_1) = m(F_1) < +\infty$, and the same applies to the positive measure μ. That is,

$$\lim_{n\to\infty} \mu(F_n) = \mu(F).$$

But $\mu(F_n) = m(F_n)$ for each n. It must be the case, then, that

$$\lim_{n\to\infty} \mu(F_n) = \lim_{n\to\infty} m(F_n)$$

which asserts that

$$\mu(F) = m(F).$$

We also know that $E \subset F_n$ for each n. Consequently $E \subset F = \bigcap_n F_n$. By our initial choice of coverings $\lim_{n\to\infty} m(F_n) = m(E)$. Since we have shown that

$\lim_{n\to\infty} m(F_n) = m(F)$, it follows that $m(E) = m(F)$. Since $E \subset F$, then $F = E \cup (F-E)$, and $m(F) = m(E) + m(F-E)$. Hence it follows that $m(F-E) = 0$. Let $\hat{E} = F - E$, and note that the beginning of the proof of Part 4 applies to $\hat{E}$, yielding a measurable set $\hat{F}$ such that $\hat{E} \subset \hat{F}$ and $\mu(\hat{F}) = m(\hat{F}) = m(\hat{E}) = 0$. Thus $\hat{E} \subset \hat{F}$ and $\mu(\hat{F}) = 0$. Hence $\hat{E} \varepsilon S$ and $\mu(\hat{E}) = 0$, since μ is a complete measure. That is, $F-E \varepsilon S$ and $\mu(F-E) = 0$. Hence $E \varepsilon S$, since $E = F - (F-E)$. Since $F = E \cup (F-E)$, it follows that

$$\mu(F) = \mu(E) + \mu(F-E) = \mu(E).$$

We now have

$$\mu(E) = \mu(F) = m(F) = m(E).$$

This proves that, whenever $E \varepsilon M$ and $m(E) < +\infty$, then $E \varepsilon S$ and $\mu(E) = m(E)$. This completes the proof of Step 4.

Step 5: If f is L-summable, then f is I-summable and $I(f) = \int_{R^k} f \, dm$.

Proof: Suppose, first, that f is L-summable and $f \geq 0$. Let $\{f_n\}$ be a sequence of functions such that $f_n \uparrow f$ and each f_n is summable and a simple function in the L-sense. That is,

$$f_n = \sum_{i=1}^{r_n} \alpha_{ni} \phi_{E_{ni}},$$

where the α_{ni}'s are real constants and the E_{ni}'s are L-summable sets.

Each $E_{ni} \varepsilon M$, and $m(E_{ni}) < +\infty$. By Step 4, each $E_{ni} \varepsilon S$ and

$$\mu(E_{ni}) = m(E_{ni}) < +\infty.$$

This shows that each f_n is I-summable and simple relative to μ. For $n \geq 1$, we have $f_n \varepsilon L^1$ and

$$I(f_n) = I\left(\sum_{i=1}^{r_n} \alpha_{ni}\phi_{E_{ni}}\right) = \sum_{i=1}^{r_n} \alpha_{ni} I(\phi_{E_{ni}})$$

$$= \sum_{i=1}^{r_n} \alpha_{ni}\mu(E_{ni}) = \sum_{i=1}^{r_n} \alpha_{ni} m(E_{ni}) = \int_{R^k} f_n \, dm.$$

The monotone convergence theorem yields the desired result that $f \varepsilon L^1$ and

$$I(f) = \int_{R^k} f \, dm.$$

Now suppose that f is an arbitrary L-summable function. Write $f = f^+ - f^-$. The above argument applies to each of $f^+ \geq 0$ and $f^- \geq 0$. Hence $f = f^+ - f^- \; \varepsilon \; L^1$, and

$$I(f) = I(f^+) - I(f^-) = \int_{R^k} f^+ dm - \int_{R^k} f^- dm$$

$$= \int_{R^k} (f^+ - f^-) dm = \int_{R^k} f \, dm.$$

This completes the proof of Step 5.

Steps 2 and 5 yield the proof of assertion (a). Steps 3 and 4 almost prove assertion (b). To complete the proof of (b), we now prove that if $E \varepsilon M$ and $m(E) = +\infty$, then $E \varepsilon S$ and $\mu(E) = +\infty$. Since the measure space (R^k, M, m) is σ-finite, we can write $R^k = \bigcup_{n=1}^{\infty} X_n$, such that $X_i \cap X_j = \Phi$ whenever $i \neq j$, each $X_n \varepsilon M$, and $m(X_n) < +\infty$. By Step 4, each $X_n \varepsilon S$, and $\mu(X_n) = m(X_n)$. That is, if $\{X_n\}$ is a decomposition of R^k guaranteed by the σ-finiteness of (R^k, M, m), then the *same* decomposition satisfies the σ-finiteness property of (R^k, S, μ). Now, for any $E \varepsilon M$ with $m(E) = +\infty$, we have

$$m(E) = m(E \cap R^k) = \sum_{n=1}^{\infty} m(E \cap X_n)$$

$$= \sum_{n=1}^{\infty} \mu(E \cap X_n) = \mu(E \cap R^k) = \mu(E).$$

The above calculation can be carried out since the sets $\{E \cap X_n\}_{n=1}^{\infty}$ are pairwise disjoint members of M with $m(E \cap X_n) < +\infty$. We have $E \varepsilon S$, and $\mu(E) = m(E) = +\infty$. This completes the proof of (b). Assertion (c) is a consequence of what has already been proved and the structure of the space R^k. If f is summable, the result is already assured by (a). The general result for f measurable is attained by considering the restrictions of f to X_n.

CHAPTER V

INTEGRATION ON LOCALLY COMPACT HAUSDORFF SPACES AND THE FUBINI THEOREM

We specialize our considerations now to Daniell integrals on vector lattices in which the original domain space X is a locally compact Hausdorff space, and L is the space $C_c(X)$. The main goal here is to obtain the Fubini theorem by a linear-functional approach.

<u>5.1 Lemma</u>: If $f_n \varepsilon L = C_c(X)$ and $f_n \downarrow 0$, then $f_n \downarrow 0$ uniformly.

Proof: Let $\varepsilon > 0$ be given. We define

$$C_n = \{x : x \varepsilon X,\ f_n(x) \geq \varepsilon\}.$$

Since f_n has compact support, $\overline{N(f_n)} = \overline{\{x : x\varepsilon X,\ f_n(x) \neq 0\}}$ is a compact set. Thus C_n is a closed subset of a compact set, and hence C_n is compact. Now since $f_n(x) \downarrow 0$ for all $x \varepsilon X$, we conclude that $\bigcap_n C_n = \Phi$. It is true, then, that $C_N = \Phi$ for some N, and hence $C_n = \Phi$ for all $n \geq N$. (This is an immediate consequence of Theorem 2.6 in Rudin [5] and the fact that $C_n \supset C_{n+1}$ for each positive integer n.) Thus

$$\|f_n\|_\infty \leq \varepsilon \qquad \text{for} \quad n = N, N+1, \ldots .$$

This shows that $f_n \downarrow 0$ uniformly, and the proof is complete.

Notation: We shall write L_c to denote the set of functions in L which vanish outside the set C. Recall that L^+ denotes the set of non-negative functions in L.

5.2 Lemma: A positive linear functional is bounded on L_c whenever C is compact.

Proof: Choose $g \in L^+$ such that $g \geq 1$ on C. Now if $f \in L_c$, $|f| \leq \|f\|_\infty$ since f is continuous. By multiplying inequalities, we see that $|f| \leq g\|f\|_\infty$. Hence

$$|I(f)| \leq I(|f|) \leq I(g\|f\|_\infty) = I(g) \cdot \|f\|_\infty,$$

and hence

$$\|I\| \leq I(g) \quad \text{on } L_c.$$

But $I(g) = M < +\infty$, and hence I is bounded on L_c.

5.3 Theorem: Every positive linear functional on $L = C_c(X)$ is an integral.

Proof: Let I be a positive linear functional on $C_c(X)$. We need to show that $I(f_n)\downarrow 0$ whenever $f_n \downarrow 0$ and each $f_n \in C_c(X)$. If $f_n \in C_c(X)$ and $f_n \downarrow 0$, then $\|f_n\|_\infty \downarrow 0$ by Lemma 5.1. Let C be the compact set $\overline{N(f_1)}$. Then $f_1 \in L_c$ and, furthermore, $f_n \in L_c$ for each positive integer n. By Lemma 5.2, I is bounded on L_c. If M is a bound for I on L_c, then

$$|I(f_n)| \leq M\|f_n\|_\infty, \quad n = 1,2,\ldots .$$

But $\|f_n\|_\infty \downarrow 0$, and thus $I(f_n) \downarrow 0$. Thus I is an integral. The proof is complete.

Notation: If X and Y are locally compact Hausdorff spaces, and I and J are positive linear functionals on $C_c(X)$ and $C_c(Y)$, respectively, then for any $f \varepsilon C_c(X \times Y)$, the notation $I_x f(x,y)$ indicates that y is thought of as fixed, in which case $f(x,y)$ is a function of x alone which is acted upon by the linear functional I. Similarly, we write $J_y f(x,y)$ when appropriate. (The notation $X \times Y$ is, as usual, used to denote the *cartesian product* set $\{(x,y) : x \varepsilon X,\ y \varepsilon Y\}$.)

5.4 Theorem: Let X and Y be locally compact Hausdorff spaces, and let I and J be positive linear functionals on $C_c(X)$ and $C_c(Y)$, respectively. Then for every $f \varepsilon C_c(X \times Y)$,

$$I_x(J_y f(x,y)) = J_y(I_x f(x,y)).$$

Furthermore, if we define $K(f) = I_x(J_y f(x,y))$ for each $f \varepsilon C_c(X \times Y)$, then K is an integral on $C_c(X \times Y)$.

Proof: Given $f \varepsilon C_c(X \times Y)$, let C_1 and C_2 be compact sets in X and Y, respectively, such that f vanishes off $C_1 \times C_2$, and let B_1 and B_2 be bounds for the integrals I and J on C_1 and C_2, respectively. Given $\varepsilon > 0$, there is a function of the form

$$k(x,y) = \sum_{i=1}^{n} g_i(x)\, h_i(y), \quad g_i \varepsilon L_{C_1}(X),\ h_i \varepsilon L_{C_2}(Y)$$

such that $\|f-k\|_\infty < \varepsilon$ for, by the Stone-Weierstrass theorem, the algebra of such functions k is dense in $L_{C_1 \times C_2}$. (See, for example, Loomis [3], page 9.) Then

$$\begin{aligned}
\left|J_y f(x,y) - \sum_{i=1}^{n} J(h_i) g_i(x)\right| &= \left|J_y f(x,y) - J_y\Big(\sum_{i=1}^{n} h_i(y) g_i(x)\Big)\right| \\
&= \left|J_y\big(f(x,y) - k(x,y)\big)\right| \\
&\le \|J\| \cdot \|f-k\|_\infty \\
&< B_2 \, \varepsilon \, .
\end{aligned}$$

This shows that $J_y f(x,y)$ is the uniform limit of a sequence of continuous functions of x. Hence $J_y f(x,y)$ is a continuous function of x. Furthermore,

$$\begin{aligned}
\left|I_x J_y f(x,y) - \sum_{i=1}^{n} I(g_i) J(h_i)\right| &= \left|I_x J_y f(x,y) - I_x \sum_{i=1}^{n} J(h_i) g_i(x)\right| \\
&= \left|I_x\Big(J_y f(x,y) - \sum_{i=1}^{n} J(h_i) g_i(x)\Big)\right| \\
&\le \|I\| \cdot \Big\|J_y f(x,y) - \sum_{i=1}^{n} J(h_i) g_i(x)\Big\|_\infty \\
&< B_1 \, B_2 \, \varepsilon \, .
\end{aligned}$$

The same thing can be done when the integrals are applied in the reverse order. That is,

$$\left|J_y I_x f(x,y) - \sum_{i=1}^{n} I(g_i)J(h_i)\right| < B_1 B_2 \varepsilon.$$

Hence

$$\begin{aligned} |I_x J_y f - J_y I_x f| &= \left|I_x J_y f - \sum_{i=1}^{n} I(g_i)J(h_i) + \sum_{i=1}^{n} I(g_i)J(h_i) - J_y I_x f\right| \\ &\le \left|I_x J_y f - \sum_{i=1}^{n} I(g_i)J(h_i)\right| + \left|J_y I_x f - \sum_{i=1}^{n} I(g_i)J(h_i)\right| \\ &< 2B_1 B_2 \varepsilon. \end{aligned}$$

But this holds for any $\varepsilon > 0$. Consequently

$$I_x J_y f = J_y I_x f.$$

Now, for $K(f) = I_x J_y f$, we note that the linearity and positivity of K are consequences of those properties for I and J:

$$\begin{aligned} K(\alpha f+\beta g) = I_x J_y(\alpha f+\beta g) &= I_x(\alpha J_y f+\beta J_y g) \\ &= \alpha I_x J_y f + \beta I_x J_y g \\ &= \alpha K(f) + \beta K(g). \end{aligned}$$

If $f \geq 0$, then $J_y(f) \geq 0$, and thus $I_x[J_y(f)] \geq 0$. That is, $K(f) \geq 0$. This shows that K is a positive linear functional on $C_c(X\times Y)$ where $X \times Y$ is a locally compact Hausdorff space. By Theorem 5.3, K is an integral. This completes the proof.

5.5 Theorem. (Fubini Theorem): Let X and Y be locally compact Hausdorff spaces. Let K, I, and J be as in the preceding theorem. (Recall that $K(f) = I_xJ_yf(x,y)$ for each $f\varepsilon C_c(X\times Y)$.) Since K is an integral, we now assume K extended by the process presented in this thesis. If $f\varepsilon \mathcal{B}^+(X\times Y)$, then

(i) $f(x,y_0)\varepsilon \mathcal{B}^+(X)$ as a function of x for each fixed $y_0 \varepsilon Y$, and $I_xf(x,y)\varepsilon \mathcal{B}^+(Y)$ as a function of y;

(ii) $K(f) = J_yI_xf(x,y)$.

Proof: Let $\mathcal{F}$ denote the class of functions in $\mathcal{B}^+(X\times Y)$ for which (i) and (ii) hold. By Theorem 5.4, $\mathcal{F}$ includes $C_c^+(X\times Y) = L^+(X\times Y)$. We now show that $\mathcal{F}$ is L-monotone. Let $\{f_n\}$ be a sequence of L-bounded functions in $\mathcal{F}$ such that $f_n \uparrow f$. Clearly $f\varepsilon \mathcal{B}^+(X\times Y)$ since $\mathcal{B}^+(X\times Y)$ is a monotone class. Also, for any fixed $y_0\varepsilon Y$, $f_n(x,y_0)\varepsilon \mathcal{B}^+(X)$ for each positive integer n. Since $\mathcal{B}^+(X)$ is a monotone class, we conclude that $f(x,y_0)\varepsilon \mathcal{B}^+(X)$. Since $I_xf_n(x,y)\varepsilon \mathcal{B}^+(Y)$ for each positive integer n and since $I_xf_1(x,y) \geq 0 > -\infty$, Theorem 2.19(b) guarantees that $I_xf_n(x,y)\uparrow I_xf(x,y)$. Since $\mathcal{B}^+(Y)$ is a monotone class, it is true that $I_xf(x,y)\varepsilon \mathcal{B}^+(Y)$. Also

$$K(f) = \lim_{n\to\infty} K(f_n) = \lim_{n\to\infty} J_y(I_x f_n) = J_y(\lim_{n\to\infty} I_x f_n)$$

$$= J_y(I_x \lim_{n\to\infty} f_n) = J_y I_x f$$

by repeated application of Theorem 2.19(b). This shows that $f \varepsilon \mathcal{F}$. Similarly, if $\{f_n\}$ is a monotone decreasing sequence of L-bounded functions in $\mathcal{F}$ and $f_n \downarrow f$, then $f \varepsilon \mathcal{F}$. We conclude, then, that $\mathcal{F}$ is L-monotone. We have shown that $\mathcal{F}$ is an L-monotone class which includes $L^+(X\times Y) = C_c^+(X\times Y)$. By Theorem 2.13 $\mathcal{B}^+(X\times Y) \subset \mathcal{F}$. But $\mathcal{F}$ was defined so that $\mathcal{F} \subset \mathcal{B}^+(X\times Y)$. Hence $\mathcal{F} = \mathcal{B}^+(X\times Y)$, and the proof is complete.

As an immediate consequence of the Fubini theorem, note that for any $f \varepsilon \mathcal{B}(X\times Y)$, it is true that $|f| \varepsilon \mathcal{B}^+(X\times Y)$, and consequently the double integral of $|f|$ agrees with the iterated integrals (both orders). If these common values are finite, then the two iterated integrals of f^+ are finite and equal, as are those of f^-. Therefore

$$|I_x J_y f - J_y I_x f| = |I_x J_y (f^+ - f^-) - J_y I_x (f^+ - f^-)|$$

$$= |I_x J_y f^+ - I_x J_y f^- - J_y I_x f^+ + J_y I_x f^-|$$

$$= |(I_x J_y f^+ - J_y I_x f^+) + (J_y I_x f^- - I_x J_y f^-)|$$

$$= 0.$$

Hence $I_x J_y f = J_y I_x f$, whenever $f \varepsilon \mathcal{B}(X \times Y)$ and $I_x J_y(|f|) < +\infty$. In other words, "the order of integration may be reversed" for a function $f \varepsilon \mathcal{B}(X \times Y)$ whenever one of the iterated integrals of $|f|$ is finite. (Of course, if $f \varepsilon \mathcal{B}^+(X \times Y)$, then the order of integration may always be reversed.) Consider the following examples involving double Riemann integrals which are of interest in connection with the Fubini theorem.

5.6 Example: Evaluate $\int_0^1 \int_y^1 \sin(x^2)\, dx\, dy$.

The integrand is continuous, and hence is a Baire function. Since the integrand is not of constant sign on the triangle in R^2 with vertices at (0,0), (1,0), and (1,1), the Fubini theorem cannot be applied as stated. However,

$$\int_0^1 \int_y^1 |\sin(x^2)| dx\, dy \le \int_0^1 \int_y^1 1\, dx\, dy = \frac{1}{2} < +\infty.$$

Thus the integrand is summable, and hence the preceding remarks apply to justify a change in the order of integration. Therefore

$$\int_0^1 \int_y^1 \sin(x^2) dx\, dy = \int_0^1 \int_0^x \sin(x^2) dy\, dx = \int_0^1 x \sin(x^2) dx$$

$$= \left[-\frac{\cos(x^2)}{2} \right]_0^1 = \frac{1}{2}(1-\cos 1).$$

5.7 Example: $\int_0^1 \int_0^1 \frac{x^2 - y^2}{(x^2+y^2)^2}\, dx\, dy$

In order to evaluate the iterated integrals, we make use of the trigonometric substitution $x = y \tan\theta$ which yields

$$\int \frac{x^2 - y^2}{(x^2+y^2)^2}\, dx = \int \frac{y^2(\tan^2\theta-1)}{y^4\sec^4\theta}\, y \sec^2\theta \; d\theta = \int \frac{\sec^2\theta-2}{y \sec^2\theta}\, d\theta$$

$$= \frac{1}{y}\int(1-2\cos^2\theta)d\theta = -\frac{1}{y}\int \cos 2\theta \; d\theta$$

$$= -\frac{\sin 2\theta}{2y} + C = -\frac{\sin\theta \cos\theta}{y} + C = \frac{-x}{x^2 + y^2} + C.$$

Thus

$$\int_0^1 \frac{x^2 - y^2}{(x^2+y^2)^2}\, dx = \left.\frac{-x}{x^2 + y^2}\right|_{x=0}^{x=1} = \frac{-1}{1 + y^2} .$$

Therefore

$$\int_0^1 \left(\int_0^1 \frac{x^2 - y^2}{(x^2+y^2)^2}\, dx\right)dy = -\int_0^1 \frac{dy}{1 + y^2} = -\arctan y\Big|_0^1 = -\frac{\pi}{4} .$$

By making the substitution $y = x \tan\theta$ and proceeding as above, or by observing the anti-symmetry of the integrand about the line $x = y$, one easily sees that

$$\int_0^1 \left(\int_0^1 \frac{x^2 - y^2}{(x^2+y^2)^2}\, dy\right)dx = \frac{\pi}{4} .$$

We have exhibited a Baire function for which the iterated integrals are finite and unequal. Note that the integrand was, and had to be, of non-constant sign. In light of the remarks following the Fubini theorem, we conclude that the integrand is not summable in the given product space, R^2.

5.8 Example: $\int_{-1}^{1}\int_{-1}^{1} \frac{xy}{(x^2+y^2)^2}\, dx\, dy$

This example shows that the summability of the integrand is not necessary (only sufficient) for the existence of finite, equal iterated integrals. Let f be defined on the square S : $|x| \le 1$, $|y| \le 1$ in R^2 by $f(0,0) = 0$, and

$$f(x,y) = \frac{xy}{(x^2+y^2)^2} \quad \text{for } (x,y) \neq (0,0).$$

For each fixed y, f(x,y) is continuous as a function of x and is an odd function, and hence

$$\int_{-1}^{1} f(x,y)dx = 0.$$

Similarly

$$\int_{-1}^{1} f(x,y)dy = 0$$

for each fixed x. Consequently

$$\int_{-1}^{1}\left(\int_{-1}^{1} f(x,y)dx\right)dy = 0 = \int_{-1}^{1}\left(\int_{-1}^{1} f(x,y)dy\right)dx.$$

We show, however, that f is not Lebesgue summable over S, i.e., that $f \notin L^1(S)$ with respect to Lebesgue measure in R^2. Since f is measurable (since f is continuous except at the origin), it suffices to show that the integral of $|f|$ over S is not finite. Let S_1 be the unit square $0 \leq x \leq 1$, $0 \leq y \leq 1$. The Fubini theorem for non-negative measurable functions guarantees that

$$\int_{S_1} f = \int_0^1 \left(\int_0^1 f(x,y)dx\right)dy.$$

A simple calculation shows that if $0 < y \leq 1$ then

$$\int_0^1 f(x,y)dx = -\frac{y}{2(y^2+1)} + \frac{1}{2y}.$$

It follows at once that

$$\int_0^1 \left(\int_0^1 f(x,y)dx\right)dy = +\infty.$$

Since

$$\int_S |f| \geq \int_{S_1} f = +\infty.$$

it follows that $f \notin L^1(S)$.

BIBLIOGRAPHY

1. Daniell, P. J., *A General Form of Integral,* Annals of Mathematics, vol. 19 (1917), pp. 279-294.

2. Heider, L. J., and J. E. Simpson, *Theoretical Analysis,* W. B. Saunders Co., Philadelphia (1967).

3. Loomis, L. H., *An Introduction to Abstract Harmonic Analysis,* D. Van Nostrand Co., Inc., Princeton, N. J. (1953).

4. Rudin, W., *Principles of Mathematical Analysis* (Second edition), McGraw-Hill, Inc., New York (1964).

5. Rudin, W., *Real and Complex Analysis,* McGraw-Hill, Inc., New York (1966).

6. Shilov, G. E., and B. L. Gurevich, *Integral, Measure and Derivative: A Unified Approach,* Prentice-Hall, Inc., Englewood Cliffs, N. J. (1966).

7. Stone, M. H., *Notes on Integration: II,* National Academy of Sciences Proceedings, vol. 34, pp. 447-455.

8. Taylor, A. E., *General Theory of Functions and Integration,* Blaisdell Publishing Co., New York (1965).